

# **The IRS Research Bulletin**

Publication 1500

1999

## **Internal Revenue Service**

### **Wayne Thomas**

Assistant Commissioner (Research and Statistics of Income)

### **Howard Oglesby**

Director, Office of Research

### **Michelle Kaplan**

Director, Office of Planning and Finance

### **Russell Geiman**

Chief, Projections and Compliance Studies Group

### **Javier Framinan**

### **Terry Manzi**

### **Melissa Kovalick**

Editors

### **Stuart Simpson**

Production Assistant

### **Product Available Electronically**

This publication is available on the IRS's Digital Daily Web site. Non-IRS users can reach that site at [www.irs.gov](http://www.irs.gov). For IRS employees, that site can be accessed at [www.irs.ustreas.gov](http://www.irs.ustreas.gov). Select the "Tax Stats" option, and then the "IRS Research Bulletin" option (under the *Statistical Publications* heading).

We welcome user feedback on this publication. Please direct questions and comments to Terry Manzi at (202) 874-1083 or Javier Framinan at (202) 874-1104 of the Projections and Compliance Studies Group, or by writing

Internal Revenue Service  
Research and Statistics of Income OP:RS  
Attn.: Projections and Compliance Studies Pub. 1500 Staff  
1111 Constitution Ave., N.W.  
Washington, D.C. 20224

We would like to thank the following individuals for their assistance in the review of this publication.

Elinor Convery  
Dennis Cox  
Charles Bennett  
Ken Hubenak  
Jeff Butler

# **The IRS Research Bulletin**

Publication 1500

1999

Suggested Citation:

United States Department of the Treasury  
Internal Revenue Service  
The IRS Research Bulletin  
Publication 1500 (Rev. 11-99)

This 1999 edition of the *IRS Research Bulletin* (Publication 1500) updates a traditional publication of the IRS Research organization. The *Bulletin* provides IRS executives, managers and staff, as well as interested external stakeholders, insights into significant trends and major IRS research findings impacting Federal tax administration. A lot has changed in the research area of IRS since our last 1995/1996 update of the *Bulletin*. This includes the further maturation of our District Office Research and Analysis (DORA) capabilities, and the implementation of a more comprehensive and systematic approach to planning and allocating resources among the Service's operational functions (based on IRS research results). In addition, even greater changes are just around the corner with the phased-implementation of the IRS modernization concepts, including the movement to the four new operating divisions. Despite these changes, and indeed because of them, this release of the *Bulletin* is perhaps more timely than ever.

IRS's vision for the future is supported by a new mission statement, a new balanced measures framework, and a renewed commitment to providing top quality service to taxpayers. This publication brings together extensive statistical information and other analytical materials which tie directly to this vision, including such IRS strategic goals as making it easier to file; providing prompt, helpful treatment to taxpayers in balance due situations; improving overall compliance; and increasing employee satisfaction. For example, the "trends" section contains a number of thought-provoking items relating to electronic commerce and e-file, customer satisfaction, and developments in the U.S. labor market, including some employee preferences.

In addition, among the research articles are two covering innovative taxpayer treatment programs designed to address noncompliance in a more "preemptive" and less intrusive fashion in the areas of self-employment taxes and duplicate claiming of dependents. A third article illustrates a new "risk-based" statistical approach to potentially screen accounts receivable cases at the point of initial assessment, thereby allowing for less forceful IRS contacts for those likely to pay, and more accelerated processing for those not likely to pay. A fourth presents the first systematic estimate of the tax gap in the estate tax area. Other articles cover: IRS efforts to provide taxpayers with needed forms in a more effective manner; an improved methodology for selecting large corporations for examination; a statistical approach for deriving fewer, yet more comprehensive, comparative measures of compliance; and a summary of the major IRS research findings in the area of electronic filing.

This update of the *Bulletin* is timely from another perspective, as well. While the modernization efforts move IRS toward new structures and new business practices, the first order of business quite often remains making sure there is a firm understanding of the past. The Research Abstracts and the indices to prior research, also contained in this publication, provide a valuable reference guide to help answer the question "has any research ever been done on ...?"

The *Bulletin* is by no means the "last word" on IRS research. Knowledge acquisition is an ongoing, cumulative process, and some of the materials presented in this edition were prepared a year or so ago. But hopefully you will find it to be the "first word," i.e., the place you start when seeking out IRS research findings, and a vehicle to identify IRS staff and/or organizations that potentially could provide you with additional insights.

Wayne Thomas  
Assistant Commissioner  
(Research and Statistics of Income)

# IRS Research Bulletin 1999 Edition

## Trends

---

1

## Articles

---

13

Alternative Treatment for Self – Employment Tax Inventory

*Kay Anderson and Dan Beckerle*

24

Duplicate Use of Dependent and Qualifying Children Social Security Numbers

*Ivette Alamo-Tirado and Robert Holmes*

31

High – Range Corporation Return Workload Selection System Development

*James A. Wilhelm*

39

Predicting Estate Tax Filings and Taxable Gifts

*Jonathan Feinstein and Chih-Chin Ho*

46

Payment Dynamics of Individual Accounts Receivable and a New Look at Risk

*Jeff Butler*

60

A Case Study of the Bank, Post Office and Library Program

*Denise York Young and Erika D. Alexander*

65

Using Data Reduction Techniques to Analyze Baseline Profiles

*Larry May and Anne Steuer*

75

Review of the IRS's Individual Return Electronic Filing and Related Research

*Javier Framinan*

## Abstracts

---

### Reporting Compliance

89

Inadequate Compensation to 1120S Corporate Officers Study

*Bruce Korbesmeyer*

90

Can Demographic Trends Predict Taxpayer Noncompliance?

*Kim M. Bloomquist*

91

Pre-payment Position and Income Tax Noncompliance

*Peter D. Adelsheim, Ph.D.*

92

Farm Labor Contractors Compliance with the Income Tax Laws

*James L. Zanetti and Marlene Le*

93

Erroneous PBA/PIA Profile

*Stan Griffin*

94

Business Profitability and Income Tax Compliance

*Kim M. Bloomquist*

95

Tax Table Clustering: Empirical Evidence of Noncompliance?

*Peter D. Adelsheim, Ph.D.*

- 96 Empirical Goodness of Fit Scoring System (Applied to Tax Preparer Due Diligence)  
*Curt Hopkins*
- 97 Processing Year 1998 Criminal Investigation Questionable Refund  
Formula Development  
*James A. Wilhelm*
- 98 Roadmap to the Future: Using Groupware as an Analytical  
Tool to Rank Emerging Compliance Issues  
*Deborah J. Meyers, Ph.D.*

### **Payment Compliance**

- 99 Closed Case Analysis to Support Collection of Agreed Deficiencies  
*Ronald L. Edgerton*
- 100 Final Study Report: Self Employed Real Estate Agents  
*Rick Denesha*
- 101 Subchapter C Corporations with Low Retained Earnings  
*Michael Hill and Richard Denesha*
- 102 Estimated Collections on FY 1998 Accounts Receivable Dollar Inventory  
*Todd Headrick*
- 103 Decline in BMF TDA Inventory in Los Angeles District  
*Scott Mendelson*
- 104 Business Licensing as a Nonenforcement Approach to Increasing  
Tax Compliance in the Liquor Industry  
*Nancy Richman and Scott Mendelson*
- 105 Analysis of Toll-Free Telephone Demand  
*Jeff Butler*
- 106 Taxpayer Service Walk-in Study  
*Curtis R. Darling*

### **Customer Service**

- 107 Selected Estimates of Returns with Form 1099 Information  
*Terry Manzi*
- 108 Alternative Signature Methods for Filing Individual Tax Returns  
*T. Scott Shutt, Robert A. Kerr and Dennis L. Raup*
- 109 Focus Group Report: On-Line Filing Program  
*Dru DeLong and David Browne*

## Statistical Tables

---

111	Table Notes
113	Table 1 Return and Economic/Demographic Data by IRS Regions and States, 1991 – 2005
132	Table 2 Return and Economic/Demographic Data by IRS Regions and Districts, 1991 – 2005
145	Table 3 Returns, FTDs, Withholding/Information Documents, and Economic/Demographic Data for the United States and IRS Centers, 1991 - 2005

## Indexes

---

151	Index of Past and Current Articles
157	Index of Past and Current Abstracts

**The views expressed in this publication represent the opinions and conclusions of the authors. They do not necessarily represent the position of the Internal Revenue Service**

## INTRODUCTION

Leaders in the Internal Revenue Service (IRS) long have been aware of the need to pay attention to the larger economic and societal trends outside the organization. The *Trends* section of *The IRS Research Bulletin* is one tool designed to help IRS management and staff with this important responsibility.

It is obvious that “external trends” impact federal tax administration and the achievement of our organizational goals. These include the types of new services the IRS needs to provide customers, the recruitment and retention of IRS’s own employees, the emergence of new compliance issues, and the opportunities for improved IRS products and services created by new technologies and other developments—to name just a few.

The *Trends* segment of the *Bulletin* is intended as a thought-provoking synopsis of some important trends impacting tax administration. On one level, the trends hopefully provide bits of useful information to IRS employees, such as by helping explain developments in the labor market and industry, or by highlighting some similarity between the IRS and the private sector. However, on another level, we also hope these trends are a catalyst for further action by IRS readers. For example, a cited trend might spark an idea for a new IRS service to meet an emerging customer need. The reader then might research the matter more thoroughly and submit a formal employee suggestion or otherwise surface their idea to management. Similarly, a listed trend might suggest to the reader a possible root cause for a particular compliance problem they have noticed recently. This reader, in turn, might seek the assistance of Research or other technical staff who systematically could investigate their causal hypothesis.

To assist the reader in reviewing the following selected trends, we have grouped them by six general areas. These are:

- Internet – Electronic Commerce – Computer Use
- Credit Cards – Debt – Bankruptcies
- Labor Market
- Workforce Characteristics and Preferences
- Employee Satisfaction
- Customer Satisfaction

Within each general area, we have attempted to sequence the trend “bullets” such that related dimensions of specific developments are within proximity.

## Internet - Electronic Commerce – Computer Use

- Although the number of electronic payment transactions is on the rise, 89 percent of all payments are still cash. Of the remaining 11 percent, 90 percent are paper checks or credit cards and the remaining ten percent are other noncash transactions.  
*(CNN interactive. October 1998)*
- Forty-two percent of U.S. households pay 1.7 monthly bills electronically. Most common methods of payment are direct debits from customers' checking accounts made to utilities, lenders, or insurers. Visa USA reports that the use of debit cards has been growing by more than 50 percent a year.  
*(American Demographics. March 1998)*
- The total number of Internet users between the ages of 18 and 34 who went online in the past month is 17 million. 7.5 million were female, and the top reasons cited for using the Internet were to seek information, e-mail, and use chat areas.  
*(American Demographics. January 1999)*
- America Online reports that the gender cybergap is narrowing, with female users accounting for 52 percent of total AOL users. Four years ago, only 16 percent of AOL members were female. Since women purchase three times as much by mail or phone orders, this narrowing gender gap has significant ramifications for online retailers.  
*(CNN interactive. August 1998)*
- The majority of American students are using computers by the time they begin school. Seven out of ten children aged six to 17 have used a computer in the past 30 days; 85 percent used them in school and 50 percent used them at home. Fifteen percent look up information about possible items to purchase and another 15 percent read periodicals online.  
*(American Demographics. April 1998)*
- Senior citizens are making huge leaps in the cyberworld. More than 20 percent own a computer, and an estimated nine million persons over the age of 50 go online. Seniors also comprise the group that most heavily uses financial service web sites and online trading.  
*(Business Week. March 1999)*
- Households with annual incomes greater than \$50,000 accounted for 74 percent of online sales. Although lower-income consumers are projected to be doing more electronic purchasing, upper-income households probably still will account for 66 percent of sales by 2003.  
*(Business Week. January 1999)*
- An estimated 15 percent of all credit card enrollments will be initiated on the Internet by 2002, while less than one percent are initiated electronically today. Although more than 15 million people conducted Internet searches for new credit cards, only six million completed online applications in the past two years.  
*(American Demographics. January 1999)*

## Internet - Electronic Commerce – Computer Use (Continued)

- In a study of 700 Internet users, the average income of persons acquiring online credit cards was \$56,315, versus the \$54,833 average income of active Internet users. People who signed up for a card online were most likely white (81 percent versus 75 percent), and more likely to live on the West Coast. Of the \$6.5 billion in online purchases during the 1998 holiday season, more than 90 percent was charged to credit cards. About half of all online shoppers say they use a single card online.  
*(American Demographics. June 1999)*
- Total electronic sales to consumers (“E-Commerce”) are expected to exceed \$18 billion for 1999, representing a \$7.8 billion increase from 1998. By 2003, electronic sales may account for as much as six percent of total consumer retail spending.  
*(Business Week. January 1999)*
- Twenty percent of firms with fewer than 100 employees used the Internet in 1996; but by 1998, that number jumped to 41.2 percent. Thirty-six percent of small businesses going online plan to use the Internet to sell their products. Revenues for small companies using the Internet average \$3.79 million, while small companies that didn’t averaged \$2.72 million.  
*(CNN interactive. July 1998)*
- A recent survey of 800 CEO’s by Price-Waterhouse Coopers revealed that half of the respondents believe online companies and other non-traditional rivals will pose a threat to their businesses in the future. However, the same group also foresees many corporate opportunities on the Internet; 40 percent think more than one-tenth of revenues will come from E-commerce in five years. Seventy-five percent of the participants reported that E-commerce currently accounts for five percent of their total revenue.  
*(Business Week. February 1999)*
- Although the media has been focusing on online consumer sales, online commerce between corporations also has been increasing significantly. Business-to-business trades over the Internet were expected to total over \$15.6 billion in 1998 and should reach \$175 billion by 2000.  
*(Time Magazine. August 1998)*
- Although E-commerce is on the rise, customer satisfaction for online purchases is going down. Cited problems include merchandise availability (15 percent), shipping and handling costs (14 percent), and slow site performance (13 percent). However, 74 percent of the holiday season online shoppers were satisfied with their overall experiences, and only five percent said they would spend less for online purchases next year.  
*(CNN interactive. January 1999)*

## Internet - Electronic Commerce – Computer Use (Continued)

- Concern for privacy remains a key issue in making online purchases. Fifty-six percent of those polled by Business Week/Harris are “very concerned” their credit information will be misused by employees where they are making their purchases, or their credit information will be made available to others without their consent. Sixty-one percent said they would be more likely to use the Internet if their privacy could be protected. Concern for privacy extends to noncredit information as well. Fifty-nine percent of those surveyed claim they never register at sites which require the disclosure of personal information, and 40 percent have given false information when registering.  
*(American Demographics. February 1999)*
- According to a recent Business Week/Harris survey, 59 percent of those polled believe that the government should pass legislation regulating Internet privacy, and only 19 percent felt that individual groups should be able to determine their own standards for privacy, without government intervention.  
*(American Demographics. February 1999)*
- In 1992, it was predicted that 100 million returns would be filed electronically by the year 2000.  
*(Fortune magazine. April 1998)*

[Editor's note: This 1992 prediction arose from a cross-functional IRS task force outside the IRS Research organization. Research forecasts of electronic filings (including those on magnetic tape) for 2000 prepared in fall 1999 project 33.6 million individual tax returns, 6.0 million business tax returns, and 0.1 million other miscellaneous tax returns.]

- The IRS estimates that the error rate on paper returns is 20 percent, as opposed to 0.5 percent for electronic returns.  
*(Business Week. February 1999)*
- One out of every five taxpayers waits until the last week to mail a return, and has spent an average of nine hours and 54 minutes working on it.  
*(CNN Interactive. April 1998)*

## Credit Cards – Debt – Bankruptcies

- Despite an average 2.5 percent fee for using credit cards to pay their taxes, more and more wealthy filers are choosing to use plastic instead of checks. The reason? Credit card incentives, such as airline frequent flier miles and other rewards.  
*(CNN Interactive. April 1999)*
- After numerous bond defaults and bankruptcies in the 1980's, corporations reduced their level of debt in the 1990's. However, net new borrowing by nonfinancial companies hit a record \$343 billion in 1998, rising by more than 10 percent annually for the first time in a decade. Meanwhile, the difference between capital expenditures and cash flow grew from \$5 billion in 1994 to \$79 billion in 1998.  
*(Business Week. April 1999)*

## Credit Cards – Debt – Bankruptcies (Continued)

- Seven out of ten Americans own at least one credit card. Thirty four percent of these cardholders do not know the interest rate of the card they use most often. People over the age of 55 are least likely to know their rate, at 44 percent. Most spenders do not realize that there is a correlation between being unaware of the cost of charging purchases and the overall increase in consumer debt.  
*(American Demographics. May 1997)*
- From 1989 to 1995, the percentage of households with credit cards rose from 56 to 67 percent. At the same time, the percentage of card-holders with incomes less than \$25,000 increased from 22 to 28 percent. The average balance increased from \$1,100 to \$1,700 during this time, while the average card-holder's liquid assets declined by more than 25 percent.  
*(Business Week. April 1999)*
- The average American consumer carries more than \$5,000 in credit card debt, while over 1.35 million Americans filed for bankruptcy in 1998.  
*(Money Magazine. April 1999)*
- Fifty-five to sixty million American households have an average debt of more than \$7,000 with over \$1,000 in annual interest and fees. Those with the most debt are households with incomes at or slightly above the federal poverty income level of \$16,036 per year for a family of four.  
*(ABCNEWS.com. July 1998)*
- Fourteen million students were projected to enroll in U.S. colleges and universities in 1998. Two-thirds of college students own credit cards and carry phone cards. The estimated spending power of all college students is more than \$90 billion, with full-time, four-year enrollees spending over \$30 billion a year.  
*(American Demographics. March 1998)*
- One in nine high school students has a credit card co-signed by a parent.  
*(Business Week. February 1999)*
- The average total debt of graduating college students in 1997 was \$18,800, compared with \$8,200, in 1991. Adding to the rising debt problem is the increasing number of students who are using credit cards to pay off tuition costs. It is estimated that \$7.5 billion will be charged this year to pay for college bills. Sixty percent of those who charge collegiate expenses pay their balances in full.  
*(CNN interactive. October 1997)*
- Approximately 1.35 million consumer bankruptcies were filed in 1997, which represents a 49.9 percent increase over the number filed in 1992. As a share of all filings, consumer bankruptcies are also higher than ever before. In 1987, consumer filings represented 85.7 percent of total filings; by 1997, the share had grown to 96.1 percent.  
*(American Demographics. July 1998)*

## Credit Cards – Debt – Bankruptcies (Continued)

- Although every state in America saw an increase in consumer bankruptcy filings from 1992 to 1997, Southern states have higher consumer bankruptcy rates overall. In 1997, Tennessee had the highest rate at 9.6 filings per thousand residents. Georgia ranked second, with 8.2 filings per thousand. Nevada, the third-ranked state at 7.9 filings, was the only non-Southern state in the top five. Finishing fourth and fifth were Alabama and Mississippi, respectively.  
*(American Demographics. July 1998)*
- Bankruptcy rates are 18 percent higher in counties with one gambling facility, and 23 percent higher in counties with five or more gambling facilities. Atlantic City has a bankruptcy rate, which is 71 percent higher than any other county in New Jersey. Clark County, Nevada, where Las Vegas is located, has the highest bankruptcy rate in the state.  
*(ABCNEWS.com. August 1997)*
- With bankruptcy filings on the rise, banks must write off increasing amounts of uncollectible debt from their credit card customers. In 1996, \$17 billion was written off as uncollectible. Regional areas with the highest numbers of filings share three basic characteristics; a high divorce rate, lax rules on automobile insurance, and a large population that lacks health insurance.  
*(Fortune Magazine. March 1997)*
- According to a study by Visa International and MasterCard International, nearly 150,000 bankruptcy filers who were pardoned of their financial obligations last year under Chapter 7 plans could have repaid 64 percent of their unsecured debts. This would have resulted in a potential recovery of more than \$4 billion.  
*(ABCNEWS.com. March 1998)*
- Examining per-capita projections of spending can provide insight into the changing fortunes of consumer goods. In 1996, the average American spent \$299 on computers (1992 dollars) and is expected to spend \$2,953 by 2006. Car sales are expected to suffer; average spending for people aged 16 and over is expected to drop from \$360 to \$302. In 1996, the average person spent \$1,313 on clothing and shoes; by 2006, this figure will increase by 18 percent to \$1,549.  
*(American Demographics. August 1998)*
- Between 1990 and 1998, the percent of owner-occupied housing units in United States counties increased from 63.9 percent to 66.3 percent. The number of business starts during this period decreased from 158.9 (thousands) to 155.1; however, the industrial production index (base year of 1987) increased from 98.9 to 131.4.  
*(Resident Population of the US, Bureau of the Census. 1998 Edition)*
- In 1997, the three states with the highest homeownership rates were Minnesota, Kentucky, and Maine. States with the lowest rates of homeowners were California, New York, and Hawaii. Nationwide, 65.7 percent of the population are homeowners.  
*(1998 Statistical Abstract of the United States. Bureau of the Census.)*

## Labor Market

- In the 1970s and 1980s, the U.S. labor force grew at a pace of 2.3 percent. In the next two decades, it is projected to grow less than one percent annually.  
*(Business Week. March 1999)*
- According to the Bureau of Labor Statistics, the U.S. labor force is expected to grow more slowly between 1996 and 2006 than it did in the previous decade. Between 1986 and 1996, the growth rate was 14 percent. From 1996 to 2006, the labor is projected to grow by 149 million participants, or 11 percent.  
*(American Demographics. March 1999)*
- Between 1996 and 2006, the labor force age 45 to 64 will grow faster than the labor force of any other age group, while the labor force 25 to 34 years of age is projected to decline by almost 3 million. The labor force participation rates of women in nearly all age groups are projected to increase, while men's labor force participation rates are expected to continue to decline for all age groups under 45 years of age. The Asian-and-other labor force and Hispanic labor force are projected to increase faster than other groups -- 41 percent and 36 percent, respectively.  
*(Employment Projections, Bureau of Labor Statistics. December 1997)*
- People residing in suburban areas are more likely to be in the labor force than those who live in central cities, regardless of age, sex, race, or ethnicity. Persons age 16 and over living in the suburbs have a labor force participation rate of 69.9 percent, and an unemployment rate of 4 percent. However, central city dwellers have a 64.6 percent participation rate, and a 7.3 percent unemployment rate.  
*("Issues in Labor Statistics," Bureau of Labor Statistics. December 1998)*
- According to Woods & Poole Economics, Inc., the number of U.S. jobs is projected to increase 12.7 percent between 1998 and 2010, to almost 172 million. However, these jobs won't be distributed evenly among the states; Summit County, Utah, is expected to see a gain of 58 percent while Edgar County, Illinois may see a 10 percent loss. Overall, southern and western counties will most likely see the greatest increase in jobs, especially those in Florida, Georgia, Texas, Colorado, Alaska and New Mexico. The prospects for the East are not as great. No eastern county appears in the top-100, and only four make the top-200: Burlington and Somerset, New Jersey; Washington, Rhode Island; and Saratoga, New York.  
*(American Demographics. May 1998)*
- As time goes by, some industries are shrinking rapidly while others are expanding. Since 1970, the number of general merchandise stores has declined from 25,032 to 14,797 in 1996. There were 1,567 drive-in theaters in 1970, compared with 408 in 1996. Some of the expanding industries include carpet/upholstery cleaning, which grew from 816 in 1970 to 8,879 in 1996. During this time frame, movie production and services increased from 2,922 to 14,680, and the number of eating and drinking establishments went from 233,048 to 466,386.  
*(Southwest Economy. January/February 1999)*
- "New Economy" industries, such as software, communications, and consulting, are adding jobs to the labor market at a rate of 3.7 percent, twice as fast as the rest of the economy.  
*(Business Week. February 1999)*

## Labor Market (Continued)

- Technical jobs comprise an average of 2.8 percent of total employment nationwide. Surprisingly, Massachusetts leads the states at 7.3 percent, beating out California at 6.8 percent. However, San Jose leads the metro areas at 28 percent.  
*(Business Week. April 1999)*
- The number of jobs lost due to mergers almost doubled between 1997 and 1998, from 37,033 to 73,903. However, the number of mergers announced in 1998 reflected only a 4 percent increase from 1997. Deregulated industries, such as telecommunications and financial services, received a disproportionate share of layoffs. Throughout the 1990s, white-collar workers continued to share the bad news with blue-collar workers.  
*(American Demographics. April 1999)*
- The unemployment rate of 4.3 percent is at its lowest point in 28 years and the service sector continues to show strong job growth. However, the manufacturing sector is continuing to cut jobs; steel makers have laid off 10,000 in the past year, while textile and apparel mills have cut 110,000 slots.  
*(Business Week. February 1999)*
- In March 1998, there were approximately 21,000 unemployed persons in the mining industry. One year later, that number jumped to 32,000. Conversely, unemployment in the construction industry decreased from 593,000 in 1998 to 490,000 in March 1999.  
*(Employment Situation News Release, Bureau of Labor Statistics. April 1999)*
- Between 1988 and 1996, the proportion of managers and professionals in the working class increased four percent, to reach a total of 17 percent by 1996. Conversely, there has been a decline in the proportion of farm workers, service employees, and craft/skilled workers.  
*(American Demographics. March 1999)*
- Faced with increasing competition as patented medications expire, drug companies are adding to their marketing departments. The top 40 drug makers now deploy nearly 59,000 representatives in the United States, up from 34,000 in 1994.  
*(Business Week. May 1999)*
- The number of automotive dealerships fell from 32,000 in 1972 to 26,000 in 1996. In contrast, employment in this area has grown from 800,000 to over 1 million during the same time. Since 1980, the occupational mix in the auto arena also has been shifting; the demand for technicians has declined while the share of supervisors has grown. Salespersons and service/parts workers have shown very little change in employment patterns.  
*("Issues in Labor Statistics," Bureau of Labor Statistics. January 1999)*
- About 12.6 million people, or one in ten workers, were classified into one of four alternative employment arrangements in February 1997. Independent contractors were the largest at 8.5 million, followed by on-call workers (2 million), temporary help agency workers (1.3 million), and contract company employees (800,000).  
*(Monthly Labor Review. November 1998)*

## Labor Market (Continued)

- After adjusting for inflation, doctors' median net income has fallen 1.4 percent each year since 1993. Before the recent managed care reform that focused on how fees were negotiated, doctors' salaries sometimes climbed 10 percent annually. *(CNN Interactive. May 1999)*
- For the third straight year, Internet workers are making the largest gains in the pay arena. Office managers and secretaries are expected to see a pay hike of 4.8 percent, while information technology workers have the highest percentage increase in starting salaries among all industries, at 7.3 percent. The large increase in salaries represents the demand for IT workers, especially in the areas of finance, insurance, and real estate. *(Business Week. February 1999)*

## Workforce Characteristics and Preferences

- The median job tenure of wage and salary workers with their current employer edged down to 3.6 years in 1998. In 1996, average tenure was 3.8 years. For males, declines in tenure were apparent in almost all of the age groups. Among females, there was very little change in overall tenure. *(Monthly Labor Review. October 1998)*
- In 1980, the average male college graduate earned about one-third more than the average male high school graduate. By 1993, the gap in earnings had increased to more than 70 percent. However, this trend may be ending for three main reasons: continued low unemployment rates leave companies with no choice but to hire less skilled workers; an increasing supply of skilled workers is holding down wage growth at the top; and information technology is becoming more user-friendly, thus enabling less-educated workers to access it more easily. *(Business Week. March 1999)*
- A recent Louis Harris poll discovered the workweek has increased by 15 percent in the last 25 years, while leisure time has decreased 37 percent. Data from the Families and Work Institute indicates 13 percent of U.S. workers are holding down two jobs. This may explain why the National Institute of Occupational Safety and Health reports that at least one quarter of today's labor force feels stressed at work *(CNN Interactive. April 1999)*
- Skilled and highly educated workers are the most likely to work longer hours. Among those with managerial, professional, or technical jobs, more than 33 percent of men and 17 percent of women put in 50-hour plus weeks, compared with 20 percent of men and 7 percent of women in other occupations. A study by Jacobs and Gerson finds almost half of the workforce would prefer to work fewer hours, and more than a quarter said they would be willing to take a pay cut to make it happen. *(Business Week. April 1999)*
- The overall number of persons working at home between 1991 and 1997 did not grow dramatically; but the number of wage and salary workers doing work at home did. In 1991, 1.9 percent of at-home workers were wage and salary; by 1997 that number had increased to 3.3 percent. Nearly nine out of ten people doing work at home were in white-collar occupations. *(Population Survey, Bureau of Labor Statistics. March 1998)*

## Workforce Characteristics and Preferences (Continued)

- According to the White House Domestic Policy Advisor, Bruce Reed, 75 percent of American families are made up of two working parents. From 1985 to 1997, the number of women with children working at home increased 32 percent, from 18 to 24 million.  
*(Business Week. May 1999)*
- Men are more likely (50 percent) to prefer full-time jobs outside the home than women (19 percent). Of those who do desire to work at home, 34 percent are white-collar workers and 33 percent are dual-earner couples. Working at home is least appealing to those age 60 and over (16 percent) and African-Americans (18 percent).  
*(American Demographics. May 1998)*
- Women tend to be more conservative when it comes to selecting a workplace. Only 49 percent of men insist on working for a profitable company, while 79 percent of women do. Women also prefer to work in team-oriented environments and in midsize companies with 100 to 500 employees.  
*(Business Week. March 1999)*
- Although the share of workers who were union members fell from 14.1 percent in 1997 to 13.9 percent in 1998, the overall number of union members rose for the first time in five years. There was little change for private industry workers, but the number of union members among government employees rose from 6.7 to 6.9 million. Local governments were most likely to be unionized, followed by the federal government.  
*(Monthly Labor Review. January 1999)*
- According to a recent AFL-CIO poll of the general public that was employed in non-supervisory jobs, 44 percent said that they would vote in favor of forming a union at their workplace. An additional 20 percent were less certain, but still positive to the issue of a union, saying that it was better to join together at a work site to solve problems.  
*(American Demographics. March 1999)*
- In 1996, six in ten job changers cashed out their retirement savings, instead of rolling them over into IRA's or employee-sponsored plans. Eighty-one percent of workers with savings less than \$3,500 cashed out, while only 14 percent with savings between \$50,000 and \$100,000 followed suit.  
*(American Demographics. April 1999)*
- In May 1991, 15.1 percent of full-time wage and salary workers were on a flexible work schedule allowing them to vary their starting and ending times. In May 1997, 27.6 percent (about 25 million workers) were on such a schedule. Executives, administrators, and managers are most likely to be on a flex schedule (42.4 percent), while less than one-fourth of those employed in administrative support or services have the option to do so.  
*(Bureau of Labor Statistics News Release. March 1998)*

## Workforce Characteristics and Preferences (Continued)

- According to a 1990 to 1997 Federal Reserve Bank of St. Louis/U.S. Census study which focused on the migration of 59 metro areas with population over one million, the following are the top five most livable: Las Vegas, Nevada; Atlanta, GA; Phoenix, AZ; Austin, TX; and Raleigh-Durham, NC. The bottom five metro areas are Orange County, CA; Miami, FL; San Jose, CA; New York City, NY; and Los Angeles, CA. (*Business Week. May 1999*)

## Employee Satisfaction

- Starting in 2000, part of the compensation for top United Airlines Inc. executives will be tied to worker satisfaction as measured by an outside survey firm. Their bonus pay also will be based on customer satisfaction and on-time performance. Together, the three new criteria will account for more than half of what the top 625 UAL managers receive. (*Business Week. March 1999*)
- Chuck Knight, CEO of Emerson Electric, firmly believes sales growth leads to a better bottom line. To help get his message across, more than half of executives' annual incentive pay is linked to revenue growth, as opposed to earnings growth, ROE, and other traditional measures. (*Fortune Magazine. April 1999*)
- FedEx's "People-Service-Profit" philosophy is based on the idea that a motivated and conscientious workforce will provide professional service to customers, which in turn will ensure profits and continued corporate growth. (*FedEx Homepage – Work Culture and Diversity; September 1999*)
- IBM offers different programs to compensate employees for their contributions. Base Pay reflects long-term responsibilities and skills, whereas the Variable Pay Program is designed to reward employees based on how they, their business units, and IBM perform against annual objectives in key areas. (*IBM Homepage – Benefits Section; September 1999*)
- Continental Airlines provides employees with the opportunity to share their success through the Stock Purchase Plan and the Profit Sharing Plan. In addition, Continental offers a cash bonus to employees on a monthly basis as a reward for their dedication and teamwork when their on-time goals are met. (*Continental Airlines Homepage – Benefits and Incentives; September 1999*)
- Companies with employee stock-ownership plans have average shareholder returns of 26.1 percent; those without ESOPs have an average return of 19.2 percent. (*Business Week. May 1999*)
- Southwest Airlines allows eligible employees to participate in a Profit Sharing Plan, funded by company contributions to profit sharing accounts. Contributions are made to the accounts when Southwest meets profitability goals set each year. Employees also can share in Southwest's success by investing in their stock, which is purchased through payroll deductions at a reduced rate. (*Southwest Airlines HomePage - Employee Benefits Summary; September 1999*)

## Employee Satisfaction (Continued)

- The ownership stakes of directors and senior officers of publicly traded U.S. corporations rose from 12.9 to 21.1 percent of their company's stock from 1935 to 1995. Meanwhile, the average value of their combined holdings increased from \$18 million to \$73 million (measured in 1995 dollars).  
*(Business Week. April 1999)*
- In 1997, seven percent of managers feared for their jobs due to poor corporate results; by 1998, 24 percent expressed concern.  
*(Business Week. March 1999)*
- According to an audit by the General Accounting Office, approximately 15 percent of IRS employees who were investigated in IRS thefts from 1995 to 1997 lacked adequate background investigation checks.  
*(CNN Interactive. December 1998)*
- Of the 87 claims of rights violations against IRS agents from 1996 to 1998, none were successful. Ninety-three percent of lawsuits filed against IRS enforcement agents during these years involved motor vehicle accidents, while only three cases involved the use of search warrants by agents.  
*(CNN Interactive. April 1999)*

## Customer Satisfaction

- IRS is not the only public agency under watch; 36 states currently have annual report cards on all of their schools, and 13 of these states require the evaluations to be sent to all parents.  
*(CNN Interactive. April 1999)*
- According to a poll by Pew Research, 60 percent of people surveyed had an unfavorable opinion of the IRS. Forty-four percent gave poor ratings to Congress, and the Pentagon and Postal Service were viewed unfavorably by 19 and 11 percent, respectively.  
*(CNN Interactive. April 1998)*
- According to a Harris Poll, over 75 percent of the population surveyed believe they have been treated fairly by IRS workers. Eighty-three percent used the word "courteous" to describe Service employees they had encounters with. The majority of Americans also believe that evasion of taxes is more rampant than IRS harassment of taxpayers.  
*(Money magazine. April 1998)*
- Only 42 percent of Americans believe that they get their money's worth from federal income taxes, 48 percent say they do not, and eight percent do not know. Women are less likely to think that their tax dollars are spent wisely, at 38 percent, compared with 46 percent of males. Forty-six percent of people age 65 and above feel positive about the way tax dollars are spent, compared with 39 percent of 18-to-24-year olds. Over half of the households with incomes under \$15,000 believe that the benefits are worth the expense, while only 40 percent of households with incomes over \$65,000 agree.  
*(American Demographics. April 1997)*

## Customer Satisfaction (Continued)

- According to a four-year assessment of 15 government agencies by Syracuse University, the Internal Revenue Service ranked 12<sup>th</sup> on the basis of factors such as management of finances, human resources, information technology, capital investment, and managing for results. The IRS received a grade of "C" overall, with individual factor grades of B/C/D/NA/B, respectively. Social Security Administration ranked first, and the Federal Aviation Administration received the lowest score. *(Washington Post. February 1999)*
- The U.S. tax code began in 1913 as a 14-page law with a one page form. In 1998, the simplest form, the 1040EZ, had a 28-page instruction book. It takes the average taxpayer ten hours to complete the regular 1040. *(Fortune magazine. April 1998)*
- In the past, the IRS often has been criticized for providing poor customer service; but things are improving. This year, taxpayers got 13 million fewer busy signals and 91 percent of the toll-free calls are being answered. In 1997, callers got through only 66 percent of the time, and only 39 percent in 1995. Nationally, accuracy scores are up to 93 percent as opposed to 63 percent in 1989. IRS employees in Baltimore gave correct advice 100 percent of the time in a random test. *(CNN Interactive. April 1998)*
- In an effort to improve customer service, the IRS now is posting "special taxpayer alerts" on its Web site to describe errors and other problems, the number of people affected, where they may live, and what they can do about it. In addition, the IRS unveiled a draft proposal that would provide \$2 million in grants to organizations providing legal assistance for low-income taxpayers involved in tax disputes. *(CNN Interactive. January 1999)*
- A recent survey compiled by field Taxpayer Advocates listed the 20 most serious problems facing tax practitioners. Topping the list was complexity of the tax laws, including figuring out exemptions, filing status, and the EITC (Earned Income Tax Credit). In second place was customer service/telephone access, with most complaints focusing on inconsistent answers and inconvenient times and locations of help centers. In third place was the cost of electronic filing, followed by offer-in-compromise (OIC) program issues. And the fifth biggest problem area focused on penalties, such as inconsistency in applying criteria and the use of penalties as a negotiation tool. *(Daily Tax Report, Bureau of National Affairs. January 1999)*

## Alternative Treatment for Self - Employment Tax Inventory

By Kay Anderson and

Dan Beckerle

*The self-employment tax (SET) inventory consists of approximately 400,000 Individual Income Tax Forms 1040 with reported income which appear to be subject to self-employment tax but do not have Schedules SE attached. Kansas City Service Center personnel and Kansas-Missouri District Office Research and Analysis (DORA) staff tested an educational letter on a sample of taxpayers from the SET inventory, tracking responses and measuring the letter's impact on compliance. Over 20 percent of the taxpayers that received the educational letter voluntarily amended their returns and on average paid an additional \$260 in tax and interest due. The Kansas-Missouri DORA subsequently helped determine at which point in the processing stream the educational letter should be generated and mailed. The team also investigated ways to eliminate from the SET inventory those taxpayers not liable for self-employment tax. The success of this research lead to its full implementation as one of IRS's ten National Strategies to increase compliance using non-traditional enforcement methods.*

### Introduction

Individuals are required to pay self-employment tax on their self-employment income.<sup>1</sup> On an income tax return, self-employment income is reported on either Schedule C, Schedule F, and/or line 21 of Form 1040<sup>2</sup>. Self-employment income also should be reported on lines 1 and 2 of Schedule SE. During filing year (i.e., return processing year) 1996, approximately 19.7 million individuals filed a Schedule C or F, or otherwise reported self-employment income on their return.<sup>3</sup> In a typical year, however, more than 400,000 individual tax returns are filed with what appears to be self-employment income but without a Schedule SE attached, as required. These returns comprise the self-employment tax (SET) inventory, also referred to as the V-code inventory.<sup>4</sup>

After the Martinsburg Computing Center identifies the SET inventory, it is made available to the Service Centers as discretionary work for their Correspondence Examination (Corr Exam) programs. The SET inventory ordered by Corr Exam at each Service Center for file years 1994, 1995 and 1996 is shown in Table 1. Most Service Centers audit some portion of their ordered SET inventory, depending on their available staff and the returns' potential dollar yield.

The Kansas-Missouri (KSMO) District Office Research and Analysis (DORA) and the Kansas City Service Center (KCSC) tested the use of an educational letter as a way to bring taxpayers in the SET inventory into compliance without the use of audits. If successful, this treatment would address the non-compliance of the unworked SET inventory, as well as some of the inventory normally ordered for examination (i.e., audit).

In this report, we (the research analysts from the KSMO DORA) first explain the method used to test the educational letter and the limitations of the test. Next, we report our findings and the estimated costs and benefits of full implementation of the tested treatment. Lastly, we discuss the conclusions we drew from this research.

<sup>1</sup> United States Code, Title 26 Internal Revenue Code, Subtitle A, "Income Taxes," Chapter 2, "Tax on Self-Employment Income."

<sup>2</sup> "Other Income" is currently, and since tax year 1994 has been, reported on Line 21 of Form 1040. However, for the tax year on which the research study was conducted, 1993, and prior tax years, "Other Income" was reported on Line 22.

<sup>3</sup> Statistics of Income Division report "U.S. Total of All Returns: Selected Sources of Income, Exemptions, Deductions and Tax by Size of AGI," report symbol NO-R:S-87.

<sup>4</sup> The SET inventory consists of all returns assigned the audit code "V" by the Martinsburg Computing Center. For tax year 1993, individual income tax returns were assigned the "V" audit code if there was net positive income in excess of \$400 from either Schedule C, Schedule F, and/or Line 22 "Other Income," wages subject to FICA were less than \$57,600, and there was no Schedule SE with the return.

**Table 1. Self-Employment Tax (SET) Inventory by Service Center and Selected File Year.**

Service Center	File Year 1994		File Year 1995		File Year 1996	
	Ordered	Total	Ordered	Total	Ordered	Total
Andover	8,036	38,764	6,000	34,168	21,681	30,459
Atlanta	6,000	29,394	0	29,748	4,057	31,108
Austin	8,531	76,210	0	77,610	0	68,698
Brookhaven	0	19,433	0	21,266	3,235	22,492
Cincinnati	10,000	74,312	10,000	76,718	0	83,228
Fresno	19,998	36,364	0	32,044	7,721	33,181
Kansas City	8,000	38,551	5,000	34,918	9,096	35,711
Memphis	2,000	28,356	5,000	29,104	8,000	29,588
Ogden	10,000	45,397	200	52,652	6,000	53,951
Philadelphia	7,992	36,029	0	46,944	0	41,041
<b>Total</b>	<b>80,557</b>	<b>422,810</b>	<b>26,000</b>	<b>435,172</b>	<b>59,790</b>	<b>429,457</b>
<b>Percent</b>	<b>19.0%</b>	<b>100%</b>	<b>6.0%</b>	<b>100%</b>	<b>13.9%</b>	<b>100%</b>
<b>Extraction Date</b>	<b>12-28-94</b>		<b>12-27-95</b>		<b>12-24-96</b>	

### First Mailing – Pilot Test in the Kansas City Service Center

In 1993, the KCSC initiated a project that tested the impact on compliance of sending educational letters to taxpayers with less than a certain self-employment income threshold. The test was limited to this group (herein referred to as the “lower strata”) because the KCSC generally orders SET inventory returns with self-employment income above the threshold (herein referred to as the “upper strata”) for possible examination. A group of taxpayers from the lower strata of the SET inventory for the 1991 tax year was selected randomly to receive an educational letter explaining self-employment tax requirements. The letter encouraged taxpayers to file amended returns if they determined they were liable for the tax. The mailing included the necessary forms and instructions to make filing an amended return easier. All responses to the letter were voluntary. The educational letter proved to be effective.<sup>5</sup> Over 20 percent of the taxpayers

<sup>5</sup> “Enhancing Voluntary Compliance with the Self Employment Tax Provisions of the Internal Revenue Code.” Unpublished: Department of the Treasury, Internal Revenue Service, Kansas City Service Center, August, 1994. The differences between the 1993 study by KCSC and the 1995 joint KCSC/DORA study are discussed in the study plan for this project, “Study Plan: Issues Associated With Self-Employment Tax Compliance.” Unpublished: Department of the Treasury, Internal Revenue Service, Kansas-Missouri District Office Research & Analysis, June 17, 1996.

receiving the educational letter took some form of corrective action.

### Second Mailing – More Detailed Study

In 1996, KCSC personnel and the KSMO DORA staff conducted a second test of an educational letter mailing to the lower strata of the KCSC SET inventory. The primary objective of this study was to determine more precisely the effectiveness of an educational letter in bringing taxpayers in the lower strata SET inventory into voluntary compliance (i.e., “voluntarily” in response to an educational letter rather than a Correspondence Examination). A second objective was to quantify the amount of self-employment tax not being collected from taxpayers in the lower strata of the SET inventory. Responses to the mailing also provided information about why taxpayers did not file the Schedule SE, what percent were misclassified workers, what impact paid tax return preparers had on the SET inventory, and to what extent unclaimed but deductible expenses reduced the estimated employment taxes due.

For this project, we randomly selected a sample of 2,189 taxpayers from the 25,469 returns in the lower strata KCSC SET inventory for tax year 1993,<sup>6</sup> and mailed them one of two letters

<sup>6</sup> The unworked tax year 1993 returns were extracted on July 25, 1995, from returns processed during file year 1994.

on January 19, 1996 (see the Appendix for copies of the letters).<sup>7</sup> We included forms, publications, instructions, and a self-addressed return envelope in the correspondence to reduce burden for those taxpayers who needed to file amended returns. All responses were voluntary. A control sample of 500 taxpayers not sent the letter was drawn from the lower strata for comparison purposes.

We tracked both telephone and written responses. The educational letters included a toll-free telephone number that routed calls to a site specifically designated for the test at the KSMO District's Taxpayer Service location in the St. Louis post of duty. The Taxpayer Service Representatives (TSRs) answering calls to this phone were provided with a questionnaire to assist them in determining the caller's self-employment tax liability. The TSRs recorded each caller's identification along with their responses to the questions. All written responses received by the KCSC were forwarded in their entirety (except for original returns and payments) to the KSMO DORA, as well.

The KSMO DORA reviewed and analyzed the written and telephone responses. We then conducted a study of those taxpayers that did not respond to the letter. We assumed the majority of non-response to the mailing occurred for two main reasons--either the letter's recipients were not liable for the tax, or they were liable but chose to ignore the letter. To corroborate these assumptions, we drew a subsample of 504 taxpayers from the original sample of 2,189 that were sent the mailing, and ordered their returns for delivery from the Federal Records Center. These 504 returns were classified for audit. Taxpayers not selected for audit included those that (1) responded to the letter and were found not liable, (2) had already filed an amended return, and (3) had SET liability below audit productivity criteria. The remaining were sent to KCSC's Corr Exam unit for audit. Of the 504 returns, 173 ultimately were audited.

<sup>7</sup> If the taxpayer reported income on Line 22 of Form 1040 ("Other Income") and therefore possibly was liable for either Self-Employment Contributions Act (SECA) tax or their share of Federal Insurance Contributions Act (FICA) tax, they were sent the first letter shown in the Appendix. If a taxpayer reported income only on Schedules C and/or F, they were potentially liable for only SECA tax and were sent the second letter shown in the Appendix .

*Tracking Amended Returns.* A number of respondents indicated they would file amended tax returns; however, resources were not available to conduct Integrated Data Retrieval System (IDRS) research on the entire sample of 2,189 taxpayers sent the letter to determine how many amended returns were filed. Ultimately, we conducted IDRS research on a representative sample of 1,000 taxpayers sent the letter (including the 504 considered for audit), plus the 500 taxpayers in the control group not sent the letter.<sup>8</sup>

The analysis showed some taxpayers filed amended returns for reasons other than receipt of the educational letter. For the purpose of this study, only amended returns to tax year 1993 with a code indicating a change to self-employment tax were classified as having been affected by our educational letter.

Also, we conducted IDRS research on the control group of 500 not sent the letter to estimate the natural rate at which taxpayers file amended returns without the influence of the educational letter.

Finally, we obtained data for the KCSC SET inventory for years both preceding (tax year 1992) and succeeding (tax year 1994) the tax year 1993 under study. We matched these data against the 1993 data to get a measurement of the amount of turnover that occurs in the SET inventory.

### **Taxpayers Liable for Self-Employment Tax**

Of the 504 taxpayers drawn for the audit subsample, 217 taxpayers (43.1 percent) were found liable for self-employment tax, 200 (39.7 percent) were found not liable, and no determination could be made on the remaining 87 (17.3 percent).<sup>9</sup>

<sup>8</sup> IDRS is the computer interface with Masterfile, IRS' computer record of tax return and taxpayer information. IDRS is the system used to retrieve tax return information from IRS' computer records.

<sup>9</sup> We were unable to determine liability of 42 taxpayers because their cases were unresolved as of February 1997. (They either went from Examination Division to default status without making a payment, were still open in Exam, or had statutory notices still pending.) We did not make determinations on an additional 37 taxpayers who did not meet audit productivity criteria and thus were not examined. We were unable to make determinations on eight other taxpayers for various other reasons.

We categorized a taxpayer as liable in two circumstances: either if they filed an amended return, or as determined by an audit by Examination. We did not count as liable those taxpayers whose examinations closed in default status (i.e., the taxpayer did not respond to the Revenue Agent's Report or the subsequent letter IRS sent assessing additional tax), unless IDRS indicated a payment to the taxpayer's account.

*Amended Returns.* Of the 1,000 test subject taxpayers for which IDRS research was conducted, 223 (22.3 percent) filed amended returns and paid an additional \$58,094 (average of \$260 per amended return) in taxes and interest. Projecting this proportion to all recipients of the mailing, we estimate 488 ( $\pm 57$  at 95-percent confidence) taxpayers filed amended returns as a response to the educational mailing and remitted an additional \$127,000 ( $\pm \$20,000$  at 95-percent confidence) in taxes<sup>10</sup> and interest.

*Control Group.* Of the 500 taxpayers in the control group who did not receive the educational letter, only six (1.2 percent) amended their 1993 returns. Three of those amended returns had adjustment reason code 44, indicating a change to self-employment tax. Therefore, we conclude the educational letter had a significant, measurable impact on the rate at which taxpayers filed amended returns.

*Audited Returns.* Of the 173 returns that were selected for audit, Corr Exam assessed \$33,214 in taxes and interest – an average of \$192 per return. Twenty of the audits resulted in no change. As of the February 1997 conclusion of the study, several cases remained in open status and \$29,438 of the assessments from these cases had been collected.

### Estimated Revenue and Costs

Using our estimated percent of the lower strata liable for self-employment tax and the dollars collected from taxpayers in our sample, we projected the total amount of self-employment tax not being collected from the lower strata nationwide. We used the file year 1995 SET

<sup>10</sup> Taxpayers are allowed an adjustment to income in the amount of one-half of their self-employment tax. This reduction in taxable income results in reduction in income tax liability. All payments of tax reported here are self-employment taxes net of any reduction in income taxes.

inventory levels (344,973 lower strata returns) and a two-sided, 95-percent confidence interval. For our subsample of 504 taxpayers considered for audit, the mean tax liability (not including interest) was \$97.45, with a standard deviation of \$151.74. Thus, we projected the uncollected tax for the total population of lower strata returns to be \$33,600,000  $\pm$  \$4,570,000.

We also projected the costs and benefits of implementing the tested treatment to the entire SET inventory nationwide. In making these projections, we estimated *minimum revenue* amounts by using the lower, one-sided, 95-percent confidence limits and *maximum costs* by using the upper, one-sided, 95-percent confidence limits. We also assumed the inventory could be reduced by filtering out taxpayers who are under the age of 18 and taxpayers who report non-passive losses on Schedule E which offset Schedule C and/or Schedule F income.

We projected the benefits of a nationwide mailing for the lower strata as described above. Although we did not test the treatment on the upper strata, we estimated the *potential* revenue that the letter would generate from this strata by assuming that, except for the amounts of self-employment income, the upper strata would respond in a manner similar to the lower strata.<sup>11</sup> The results of the estimated revenues are shown in Table 2. These estimates do not include any audit revenues or costs.

<sup>11</sup> The study team concluded that the most cost effective option for testing (in effect) the educational letter on the upper strata was to administer it to the entire upper strata inventory concurrently with treatment of the lower strata. Such treatment at a nationwide level would add only \$140,000 to the project cost while potentially adding significantly more revenue.

**Table 2. Estimated Revenue, Cost, and Return on Investment for Nationwide Mailing.**

Strata	Tax & Interest Collected	Costs to Collect	Return on Investment
Projected (Lower Strata)	\$14,100,000	\$536,000	26
Potential (Upper Strata)	\$29,500,000	\$140,000	211
Total	\$43,600,000	\$676,000	64

Even if no revenue is generated by sending the letter to the upper strata, the return on investment is 21 ( $\$14,100,000/\$676,000 = 21$ ).

**Taxpayers Not Liable for Self-Employment Tax**

We found approximately 40 percent of the lower strata SET inventory not liable, based on the 504 taxpayers considered for audit. Taxpayers provided numerous reasons for not being liable. Two explanations accounted for 36 percent of the non-liable respondents. The first involved taxpayers claiming to have erroneously filed a Schedule F (reporting profit or loss from farming on the Individual Income Tax Form 1040) rather than a Form 4835 (reporting farm rental income and expenses on the Form 1040). The second was for taxpayers under the age of 18 and employed either by their parent(s) or as a newspaper carrier. Table 3 shows the various explanations (and associated frequencies) taxpayers provided. Further description of these explanations follow.

**Table 3. Reasons for Not Being Liable for Self-Employment Tax.**

Reason for not being liable	Count	Percent
Filed Schedule F instead of Form 4835	41	20.5%
Under Age 18	31	15.5%
Included Both Spouses' Incomes	16	8.0%
Sched. E Loss Offset Sched. C or F Income	13	6.5%
Administrator/Executor	7	3.5%
Employee (Including Statutory)	7	3.5%
Prizes/Awards	7	3.5%
Minister/Religious Exemption	7	3.5%
Sale of Asset	7	3.5%
Schedule E Income	6	3.0%
One-time Activities	5	2.5%
Strike Benefits	3	1.5%
Expense Reimbursement	3	1.5%
Gambling Winnings	3	1.5%
State Retirement System	3	1.5%
Additional Expenses	3	1.5%
All Others	38	19.0%
Total	200	100%

*Schedule F Returns.* Of the 504 taxpayers in the subsample, 75 filed Schedule F, Profit or Loss from Farming, with their Individual Income Tax return (Form 1040). More than half of these taxpayers (41) were not liable for self-employment tax because they rent their farms and rental income is not subject to self-

employment tax. However, these taxpayers erroneously reported their rental income on Schedule F; this income should be reported on Form 4835, Farm Rental Income and Expenses and Schedule E. In total, 76 percent (57 of the Schedule F filers in our subsample) were not liable for self-employment tax.

*Taxpayers Under Age 18.* IRS records indicate that 28 of the 504 subsample taxpayers were under the age of 18 during the 1993 tax year. Of those, we determined 24 to be not liable, three to be liable, and one was indeterminable. Through direct communication with the taxpayer, we determined three others in the subsample, for whom there was no record of age, were not liable because they were under age 18 and employed either by their parent(s) or as a newspaper carrier.

*Spouses With Self-Employment Income Less Than \$400.* A tax return is included in the SET inventory when the combined total from all Schedules C and Schedules F included with the return exceed the self-employment income threshold of \$434 per individual.<sup>12</sup> Although the combined self-employment income reported on their returns exceeded the threshold, 16 taxpayers in the subsample (8.0 percent of those not liable) were not liable for self-employment tax because each spouse's individual self-employment income was less than \$434.

*Schedule E Losses.* Thirteen respondents (6.5 percent of those not liable) in the subsample were not liable for self-employment taxes because they had Schedule E losses which offset Schedule C and/or Schedule F income. Taxpayers should reduce their Schedule C and/or Schedule F income subject to self-employment tax by non-passive partnership losses reported on Schedule E.

*Administrator/Executor.* Seven taxpayers in the subsample indicated they had received compensation as a result of serving as either an executor of a relative's estate or as the administrator of an incapacitated relative's affairs. Since these taxpayers were not in the business of serving as an executor or administrator, they are not subject to self-employment tax.

*Statutory Employees.* Statutory employees<sup>13</sup> accounted for 3.5 percent of the subsample that

<sup>12</sup> Self-employment tax applies only to the amount of self-employment income in excess of \$400. The amount of self-employment income subject to this threshold is equal to combined Schedule C and Schedule F profit/loss reduced by 7.65 percent. Thus, the effective threshold is  $\$400/(1-.0765) = \$434.13$ .

was not liable for self-employment tax. Nationwide, approximately 65,900 workers are employed as statutory employees. Although statutory employees make up only 0.06 percent of all filers, they are disproportionately represented because they comprise 3.5 percent ( $\pm 1.6$  percent at 95-percent confidence) of the SET inventory.

*Minister/Religious Affiliation.* Seven taxpayers in the subsample were not liable for self-employment tax because they were either religious clergy exempt from this tax or were members of a recognized religious group and have personally waived their rights to Social Security benefits. Taxpayers must obtain approval from the Internal Revenue Service in order to claim this exemption from self-employment tax. Although there is a minister indicator field on the Individual Master File (IMF) to capture this exemption, it did not indicate that any of these seven had a valid exemption. (All seven provided proof of their valid exemption.)

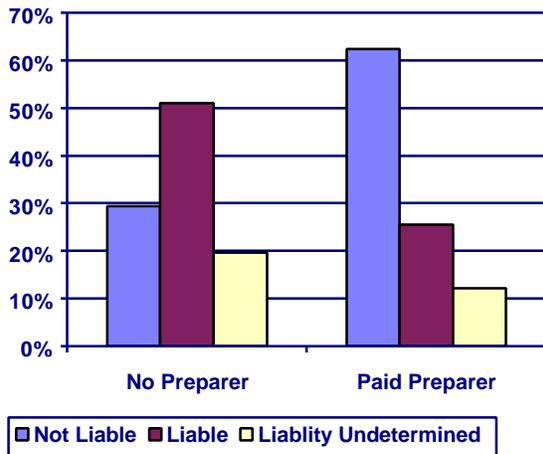
Ministers are of interest in the SET inventory because they are routinely surveyed (i.e., the returns are ordered from the records center but not examined) by Corr Exam.<sup>14</sup> If the minister indicator on the IMF could be used as a filter to eliminate these returns from the SET inventory, the cost of ordering and then surveying these returns could be saved; however, the indicator is of questionable use. Over 62 percent of the ministers in KCSC's SET inventory are in the upper strata, but the IMF indicates only 32 percent of them have an approved exemption from self-employment tax.

<sup>13</sup> Statutory employees include: (1) full-time traveling or city salespeople who solicit orders from wholesalers, restaurants, or similar establishments, on behalf of a principal; (2) full-time life insurance agents whose principal business activity is selling life insurance and/or annuities for one life insurance company; (3) agent/drivers or commission-drivers engaged in distributing meat, vegetables, bakery goods, beverages (other than milk), or laundry or dry cleaning goods; and (4) home workers performing work on material or goods furnished by the employer. Employers indicate on a Form W-2 whether a worker is classified as a statutory employee. Statutory employees report their wages, income, and allowable expenses on Schedule C, but they are not liable for self-employment tax because the employers are obligated to treat statutory employees as employees for social security tax purposes. Source: *1993 U.S. Master Tax Guide, 1941B*.

<sup>14</sup> A taxpayer is considered a minister if he/she reports Primary Business Activity Code 8771 on his/her Schedule C.

*Paid Preparers.* Respondents whose returns were completed by paid tax return preparers were far less likely to be liable for self-employment tax than respondents who prepared their own returns. Only 25.5 percent of taxpayers who used a paid preparer were determined to be liable whereas 51.0 percent of taxpayers who prepared their own returns were determined to be liable. These results are presented graphically in Figure 1.

**Figure 1. Self-Prepared Returns versus Paid Preparer Returns.**



*Inaccurately Filed Returns.* Several reasons for not being liable are grouped under the general category of inaccurately filed returns. They relate to issues associated with Schedule E income, sales of assets, state tax refund income reported on Line 22, and scholarship income (the last two of which are included in the “All Others” category in Table 3). Taxpayers in this category represent 67, or 33.5 percent, of those in the subsample who are not liable.

*One-time Activities.* The “One-time Activities” category represents taxpayers who performed some activity for a relative or friend for which they were paid; however, they were not in the business of providing this particular service to the general public. In many respects, some of the other categories listed in Table 3 also represent one-time activities, including Administrator/Executor, Prizes/Awards, Sale of Asset, and Strike Benefits. In total, 30 taxpayers (15 percent) in the subsample were not liable because of one-time activities.

**Turnover in the SET Inventory**

We evaluated the amount of turnover in the KCSC SET inventory. To do this, we matched the tax year 1993 KCSC SET inventory to the 1992 and 1994 inventories. The following shows the percentage of the 1993 SET inventory present in other years’ inventories:

SET Inventory (Year)	Percent of 1993 SET Inventory Present
1992	25.3 %
1994	22.3%
Both 1992 and 1994	10.8%
	36.7%

We concluded there was a high degree of turnover in the SET inventory with the possibility that a small core group remains in the inventory year to year.

**Conclusion**

The educational mailing proved to be an effective alternative to audits as well as an effective way to improve compliance among taxpayers in the SET inventory. Many of the taxpayers liable for self-employment tax indicated they were unaware of their liability. Once made aware of their responsibility, many filed amended returns with full payments.

The amount of tax not being collected from the lower strata of the SET inventory is significant. We project it to be \$33,500,000 (± \$4,570,000 at 95-percent confidence). Furthermore, much of the upper strata is not audited in any given year, thus likely adding millions of dollars to self-employment taxes not collected. However, limiting our revenue estimates to just the lower strata, we estimate the educational mailing will generate at least \$14 million in additional tax receipts.

In addition to responses from liable taxpayers, responses from those not liable provided valuable information. This information indicated ways to reduce the SET inventory (i.e., by eliminating those not liable), as well as to handle it in a more cost-effective manner.

### From Research to Operational Programs

Based on the effectiveness of the test, IRS adopted implementation of the educational letter treatment to the entire SET inventory as a National Strategy in Fiscal 1998. The National Strategies are a key component of IRS's Operations Plan, the multi-year, multi-functional planning document for the Chief Operations Officer area. Fiscal 1998 served as a transition year where the letter was administered in a mass mailing. For Fiscal 1999 and subsequent years, mailing of the letter will be accomplished through a computer generated notice, reducing the time between filing of the original return and receipt of the letter to less than four weeks.

Adoption of the strategy also includes identification and removal from the inventory taxpayers not liable for self-employment tax, review of the tax return processing procedures concerning self-employment tax, and possibly consideration of tax form changes to improve filing accuracy. Preliminary indications suggest the inventory can be reduced by as much as 25 percent (62 percent of those not liable) by using a more accurate selection process. Some tax return processing changes have been made, and tax form changes are under consideration.

---

#### About the Author(s):

**Kay Anderson** is an economist in the Kansas-Missouri District Office of Research and Analysis. She received her M.A. degree in economics in 1988 from Northern Illinois University. She has been with the IRS since 1995.

**Dan Beckerle** is an operations research analyst Team Leader in the Kansas-Missouri District Office of Research and Analysis. He received his B.S. in Mining Engineering in 1979 from the University of Missouri – Rolla and his M.B.A. in 1990 from the University of Houston. He has been with the IRS since 1982.

**Appendix A**

**Educational Letters**

ALTERNATIVE TREATMENT FOR SELF-EMPLOYMENT TAX INVENTORY

**Kansas City, MO 64999**

Mail Stop 4200

**Person to Contact:**

CP:RAD:PIU

**Telephone Number:**

District Office Toll Free  
Number: 1-800-736-5271

**Date:**

Taxpayer Identification  
Number:

Dear Taxpayer:

We have analyzed the information you provided on your individual income tax return for 1993. You received and reported income which could be subject to self-employment tax or your share of FICA taxes not withheld from your wages. For example, Schedule C or F net profit or certain Form 1099 income is subject to either self-employment or FICA tax.

Enclosed is a copy of Schedule SE and instructions for self-employment tax, and a copy of Employee Procedures and Employee FICA Worksheet for your share of FICA taxes to assist you in determining if you are liable for either of these taxes.

If you determine that you owe self-employment or your share of FICA tax for 1993 or 1994 please complete a separate Form 1040X, Amended US Individual Tax Return, for each year that you owe tax. The computation schedules (Schedule SE for Self-employed individuals or Employee FICA Worksheet for employees) should be attached to the amended return. If you owe self-employment or FICA tax for 1995 you should include a completed Schedule SE or Employee FICA Worksheet with your original return. Please mail completed forms and any written correspondence to the IRS in the enclosed self-addressed envelope.

If you have any questions concerning this information or if you need help in any way please call the toll free number listed above. You may also contact your local Internal Revenue Service office in person.

Sincerely,

*(signed) Nancy L. Jones*  
Nancy L. Jones  
Chief, Examination Branch  
Kansas City Service Center

**Enclosures:**

Return envelope  
Copy of this letter  
Employee Procedures  
Employee FICA Worksheet  
Form SS-8  
Form 1040X  
Schedule SE and instructions

**Kansas City, MO 64999**

Mail Stop: 4200

**Person to Contact:**

CP:RAD:PIU

**Telephone Number:**

District Office Toll Free

Number: 1-800-736-5271

**Date:**

Taxpayer Identification

Number:

Dear Taxpayer:

We have analyzed the information you provided on your individual income tax return for 1993. You received and reported income which could be subject to self-employment tax or your share of FICA taxes not withheld from your wages. For example, Schedule C or F net profit or certain Form 1099 income may be subject to self-employment tax if the total is \$434.00 or greater.

Enclosed is an excerpt from Publication 533, "Self-Employment Tax", to assist you in deciding whether you are liable for the self-employment tax.

If you determine that you owe self-employment tax for 1993 or 1994 please complete a separate Form 1040X, Amended US Individual Tax Return, for each year. Forms for 1993 are enclosed. If you owe self-employment tax for 1995 you should include a completed Schedule SE with your original return. Please mail completed forms and any written correspondence to the IRS in the enclosed self-addressed envelope.

If you have any questions concerning this information or if you need help in any way please call the toll free number listed above. You may also contact your local Internal Revenue Service office in person.

Sincerely,

*(signed) Nancy L. Jones*

Nancy L. Jones

Chief, Examination Branch

Kansas City Service Center

Enclosures:

Return envelope

Copy of this letter

Excerpt from Publication 533

Form 1040X

Schedule SE and instructions

# **Duplicate Use of Dependent and Qualifying Children Social Security Numbers**

By Ivette Alamo-Tirado and  
Robert Holmes

*Building upon prior research, including Internal Audit findings of an increasing trend in the use of duplicate Social Security Numbers (SSNs), the North Florida District Office of Research and Analysis (NFL-DORA) profiled this population and tested alternative treatments to improve compliance. Our analysis of the tax year (TY) 1995 duplicate SSN population revealed it is composed largely of situations involving two returns claiming one duplicated SSN. In particular, around 3.2 million individual returns contained a duplicate SSN in TY 1995. This reflected around 1.8 million duplicated SSNs, of which 98.6 percent were claimed on only two returns. The most frequent duplication (in 46.3 percent of the 3.2 million returns) occurred where one return claimed a dependent and that dependent also filed a return and claimed a self-exemption. To address the problem of duplicate SSN usage, we tested two major alternative treatments designed to improve compliance, one involving a service center correspondence examination and the other involving a "soft" (i.e., educational) notice. Both treatments proved successful in improving subsequent compliance, reducing instances of duplicate SSN usage in the subsequent filing year by over 15 percentage points in the tests conducted. However, the finding of a relatively low year-to-year "repeater" rate among the population served to dampen some of the potential effectiveness of the treatments. Still, the research revealed insights into new approaches to better target this noncompliant market segment, and has led to several operational program initiatives. This project was adopted as one of the National Strategies within the multi-year Operations Plan for the Chief Operations Officer area and is designed to improve compliance through innovative approaches.*

## **Introduction**

Only one return legally can claim an exemption or the Earned Income Tax Credit (EITC) for a particular person. Duplicate claims for the same person result in substantial revenue loss to the government. By mid 1996, the IRS processed 3.2 million tax year (TY) 1995 U.S. Individual Income Tax Returns that contained a duplicate claim for a valid Social Security Number (SSN) as a dependent, qualifying child for EITC, and/or a dependent filing a return and claiming him/herself.

The IRS checks the validity of an SSN during return processing by confirming the Social Security Administration issued one to the taxpayer, dependent, or qualifying child. Presumably, half the population of identified duplicate SSN users is entitled to claim the exemption and/or qualifying child. With minor exception, however, data do not exist internally to the IRS to determine who is entitled and who is not. Also, the IRS cannot afford to spend significant enforcement resources to bring this entire population into compliance.

As a result, the North Florida District Office Research and Analysis (NFL – DORA) undertook this research project to profile the market segment of returns involving duplicate use of SSNs. We (NFL – DORA staff) wanted to learn more about this population, and measure the effectiveness and efficiency of "soft notices" (i.e., educational-type letters) versus correspondence examinations (i.e., audits) in improving future compliance and motivating taxpayers to file amended returns. This article presents the research results relating to the TY 1995 valid duplicate SSN population of 3.2 million returns.

## **Profile Highlights**

### Types of SSN Duplication

We identified the duplicate SSN population based on three duplicate SSN conditions:

1. duplicate EITC - two or more returns claiming the same qualifying child
2. duplicate dependent - two or more returns claiming the same dependent

DUPLICATE USE OF DEPENDENT AND QUALIFYING CHILDREN SOCIAL SECURITY NUMBERS

- 3. dependent/primary - a return claiming a dependent and that dependent filing a return and claiming a self-exemption.

Many returns in the population involved combinations of the three conditions, as the following table illustrates.

**Table 1. Number and Percentage of TY 1995 Returns by Type of Duplicate Condition**

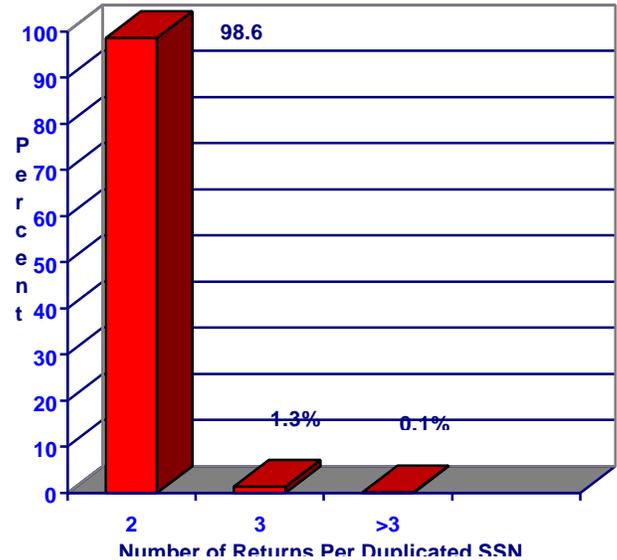
DUPLICATE TYPE	NUMBER	%
EITC Only	50,541	1.6%
EITC & Dependent	434,450	13.5%
Dependent Only	1,170,025	36.4%
Dependent & Dependent/Primary	55,878	1.7%
Dependent/Primary Only	1,488,953	46.3%
Other	18,666	0.5%
<b>Total</b>	<b>3,218,513</b>	<b>100%</b>

Only 1.6 percent of the returns in the population had just a duplicate EITC condition; however, 13.5 percent contained both a duplicate EITC and duplicate dependent condition. This situation could have involved the same SSN claimed for both EITC and dependent, or could have involved two or more different SSNs. Duplicate dependents accounted for over 50 percent of the population when considering multiple conditions (i.e., duplicate dependent and EITC, duplicate dependent and dependent filing as primary). A dependent on one return, filing their own return and claiming the self-exemption, represents a significant portion (46.3 percent) of the population.

Number of Returns Involved

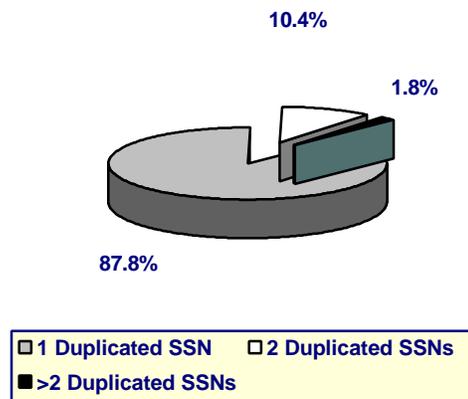
There are 1,824,108 duplicated SSNs in the TY 1995 population. Of these 98.6 percent were claimed on only two returns. Chart 1 illustrates these data.

**Chart 1. Percent of Returns by Number of Returns Per Duplicated SSN**



Of the 3,218,513 returns in the population, 87.8 percent involves only one duplicated SSN and 10.4 percent involves two. Only 1.8 percent involves more than two duplicated SSNs. Chart 2 illustrates these findings.

**Chart 2. Percentage of Returns by Number of Duplicated SSNs Per Return**



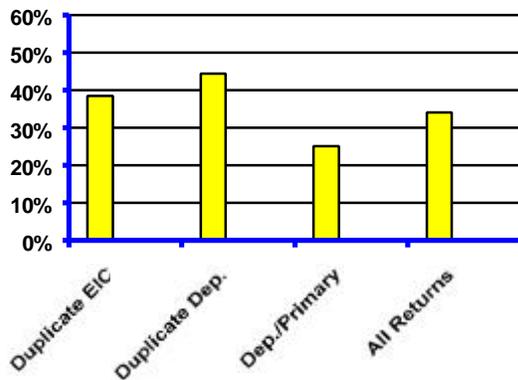
## DUPLICATE USE OF DEPENDENT AND QUALIFYING CHILDREN SOCIAL SECURITY NUMBERS

Considering 98.6 percent of the duplicated SSNs was claimed on two returns and 87.8 percent of the returns involves only one duplicated SSN, clearly the population is composed largely of two returns claiming one duplicated SSN.

### Repeater Rate

The primary SSN on a tax return reflects the person filing (or the person for whom the return is being filed). We defined a repeater as a primary SSN in the TY 1995 duplicate SSN file that also was a primary SSN in the TY 1994 duplicate SSN file. The following chart reflects the percentage of repeaters by group:

**Chart 3. Percentage of Repeaters in Population by Group**



The repeater rate for all returns in the population is only 34.1 percent. Conversely, 65.9 percent of the taxpayers in the TY 1995 file was not in the TY 1994 population. Returns with a duplicate dependent condition had the highest repeater rate at 44.5 percent. The repeater rate for returns with a duplicate EITC condition had a repeater rate of 38.5 percent. The dependent filing as primary condition had the lowest repeater rate at only 25.1 percent.

These results indicate the duplicate SSN population turns over rapidly. Hence, this limits the effectiveness of any treatment involving one-to-one contact. For all contacts, notice or examination, only the portion made with taxpayers that would repeat a duplicate SSN claim can affect future compliance.

### **Treatment Tests**

We conducted three distinct taxpayer treatment tests. The first two involved two different “soft” (educational) notices sent to different samples of the population. Both notices informed the taxpayer of the duplicate SSN condition, pointed them to readily available information sources that explained the proper circumstances for claiming the personal exemption, dependents, etc., and advised them not to claim the duplicated SSN if not entitled. However, one of the notices also requested the filing of an amended return for TY 1995 if the taxpayer was not entitled to claim the duplicated SSN. The other notice made no mention of filing an amended return.

The third test treatment involved sending a standard IRS Examination initial contact letter to samples of the population. This initial contact letter advised taxpayers of the examination of their TY 1995 tax return and requested taxpayers verify with the IRS their dependency and earned income credit claims.

We mailed the soft notices and Examination contact letters in late November and early December 1996, before taxpayers filed their TY 1996 returns. The primary purpose of the test treatments was to determine the relative effectiveness of the various contacts in improving compliance on TY 1996 returns. Another objective was to measure the effectiveness of the particular soft notice in motivating taxpayers to file amended returns for TY 1995.

### **Methodology**

#### Linking Related Returns

To enable selection of the test and control groups and evaluation of the test results, we linked taxpayers that claimed the same duplicated SSN or SSNs to one another to form a “case.” We then assigned each case a control number. Of the 3,218,513 returns in the duplicate SSN population, we successfully linked 3,177,574, or 98.7 percent, to create 1,530,230 cases.

#### Selection of Test and Control Groups

The duplicate SSN population of returns successfully linked then was segmented into six groups for sample selection. First, we split the

population into three groups based on the duplicate condition(s) – i.e., duplicate EITC, duplicate dependents, and dependent/primary. Since many cases involved multiple duplicate conditions, we used the following methodology in assigning cases to a group:

- any case involving a duplicate EITC return was assigned to the EITC group,
- any remaining case with a dependent/primary condition was assigned to the dependent/primary group, and
- any case with only a duplicate dependent condition was assigned to the duplicate dependent group.

We then segmented the three groups based on whether the case was a repeater (i.e., also present in the TY 1994 duplicate SSN population) or non-repeater, resulting in six market segments from which the test and control group samples were selected. We selected a random sample of approximately 550 cases for each of the test and control groups. Cases selected for the control groups received no IRS contact (relative to this research).

Conducting the Tests and Measurement of Results

As previously stated, in November/December 1996 we mailed two different soft notices to samples of the six segments of the population. We placed our return DORA address on the envelope so we could measure and respond to all correspondence received. In addition, a unique toll-free telephone number was established to route taxpayer inquiries to one Customer Service group for additional control and cost/benefit analysis purposes.

The third treatment test involved examinations of returns. However, this test included only the three segments of the repeater population. Examinations were not initiated on non-repeaters. In November 1996, we furnished identifying information of the taxpayers in three repeater segments to a Service Center Correspondence Examination Branch; and then in early December 1996 mailed all contact letters.

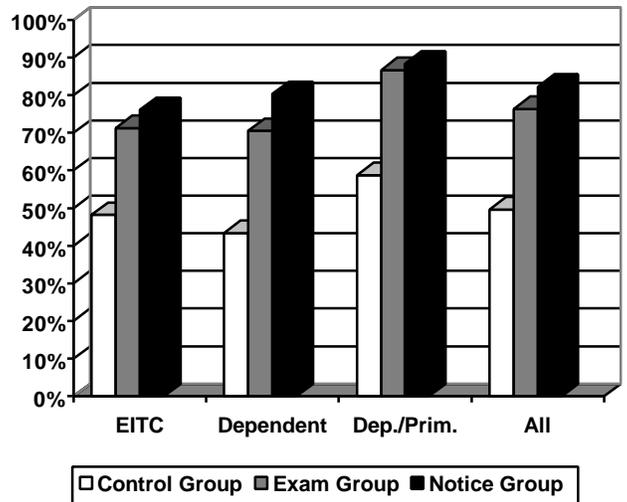
**Test Results**

We measured the compliance improvement resulting from the test treatments by searching the TY 1996 valid duplicate SSN population for the presence of the taxpayers in the 1995 test and control groups. Taxpayer presence in the TY 1996 valid duplicate SSN population required at least one other taxpayer to file a return and duplicate the SSN claim (and satisfy one of the three duplicate conditions previously described). We measured the soft notice effect on amended return filing by using IRS Master File data from the TY 1995 module of taxpayers in the test and control groups.

Effectiveness of Notices and Examinations in Improving Compliance on TY 1996 Returns

Chart 4 shows results of the repeater control group, soft notice, and Examination contact letter. Since the two soft notices were equally effective in improving compliance on TY 1996 returns, only data for the notice that requested the filing of an amended return are presented below. Again, the Examination initial contact letters were mailed to samples of the repeater population only.

**Chart 4. Percentage of TY 1995 Repeaters Not Repeating a Duplicate SSN Claim in TY 1996 by Group**



## DUPLICATE USE OF DEPENDENT AND QUALIFYING CHILDREN SOCIAL SECURITY NUMBERS

By definition, the control groups were not subjected to the treatments. Therefore, the differences between the test and control group for each the soft notice and Examination contact letter represent the treatments' effect in improving compliance for TY 1996.

Comparing the notice group with the control group, it is clear the educational notice was successful in reducing duplicate SSN claims in the subsequent year for the repeater segment. In particular, the notice improved compliance 27.9 percentage points (76.1 percent - 48.2 percent) for the EITC segment. It increased compliance 36.9 percentage points (80.3 percent - 43.4 percent) for the dependent-only segment, and 29.5 percentage points (88.3 percent - 58.8 percent) for the dependent/primary segment. Compliance increased 32.7 percentage points (82.3 percent - 49.6 percent) for repeater taxpayers, overall.

Considering the effects of the Examination contact letter, we again see a distinct reduction in the instances of duplicate SSNs, compared to the control group. The Examination contact letter increased compliance 23.1 percentage points (71.3 percent - 48.2 percent) for the EITC group, 27.1 percentage points (70.5 percent - 43.4 percent) for the dependent group, 27.9 percentage points (86.7 percent - 58.8 percent) for the dependent/primary group, and 26.7 percentage points (76.3 percent - 49.6 percent) for taxpayers overall.

Interestingly, the percent of taxpayers not repeating a duplicate SSN claim for the notice groups was higher for all groups compared to the Examination contact letter. However, the difference was "statistically significant" only for the dependent group.

Among the non-repeater population, the notice also proved effective in reducing duplicate SSN claims compared to the control group. As shown in Chart 5, the notice improved future compliance (i.e., increased the instances of no repeat duplicate SSN claims) by 16.7 percentage points for the EITC group, 12.0 percentage points for the dependent group, 5.2 percentage points for the dependent/primary group, and 8.9 percentage points for non-repeat taxpayers overall.

**Chart 5. Percentage of TY 1995 Non-Repeaters Not Repeating a Duplicate SSN Claim in TY 1996 by Group**

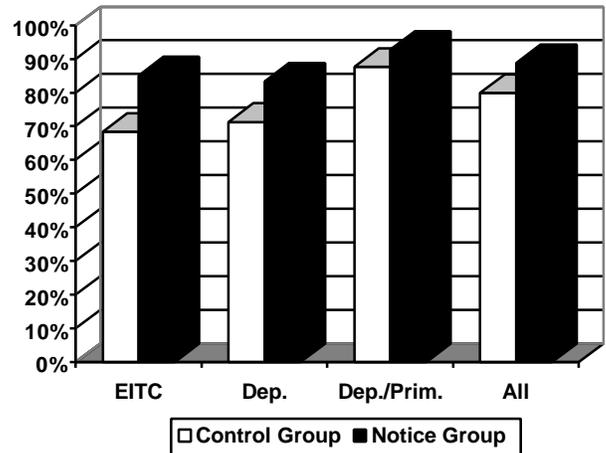
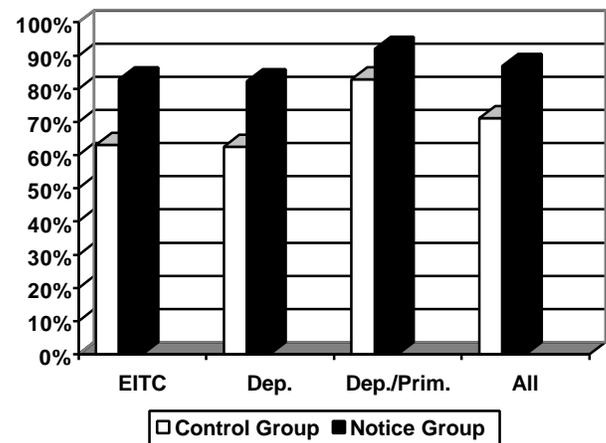


Chart 6 presents the consolidated results for all the notice and control group taxpayers in the three duplicate SSN categories. This chart weights and combines the results for repeaters and non-repeaters.

**Chart 6. Percentage of All TY 1995 Taxpayers Not Repeating a Duplicate SSN Claim in TY 1996 by Group**



Considering the entire population of repeaters and non-repeaters, the notice improved compliance by 19.7 percentage points for EITC taxpayers, 19.8 percentage points for dependent taxpayers, 9.2 percentage points for the dependent/primary taxpayers, and 15.7 percentage points for taxpayers overall.

For all notice groups, 13.1 percent repeated a duplicate SSN in TY 1996. This portion may require enforcement action to treat.

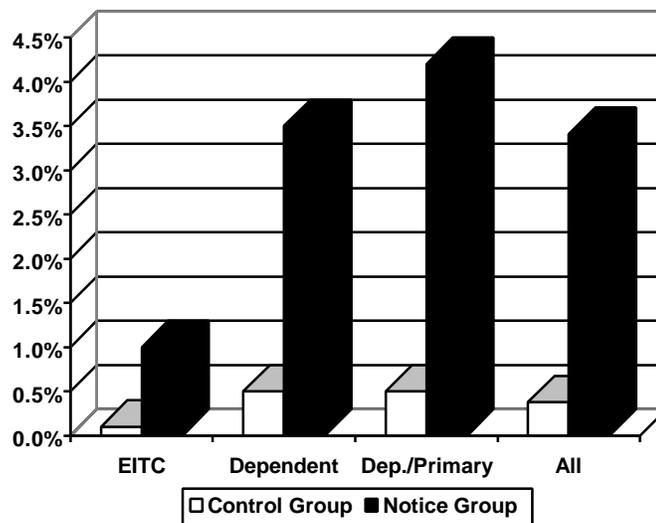
Effectiveness of Notices in Motivating Taxpayers to File Amended Returns for TY 1995

Since this project focuses on the duplicate use of an SSN, there is an inherent presumption that one of the taxpayers claiming the duplicated SSN is entitled to the claim, and one or more taxpayers are not. Since over 98 percent of the duplicated SSNs involve two taxpayers, we assume approximately half of the population is entitled to claim the duplicated SSN(s). Therefore, if all the taxpayers not entitled to claim the duplicated SSN(s) filed amended returns, the maximum amended return filings would be approximately 50 percent.

Chart 7 presents the percentage of taxpayers filing amended returns with a balance due or refund claimed. Amended returns with no change in liability were excluded due to the likelihood taxpayers filed these to notify the IRS of an SSN error. The notice did not discourage the filing of an amended return to correct SSN errors, and we received several amended returns where the taxpayer indicated they mistakenly used the wrong SSN (generally that of another child or ex-spouse).

The notice did precipitate filing of amended returns within the notice groups. Compared to the control group, instances of filing an amended return for the notice group was 0.9 percentage points higher for EITC, 3.0 percentage points higher for dependent, 3.7 percentage points higher for dependent/primary, and 3.0 percentage points higher for taxpayers overall. Still, the results indicate the notice is more likely associated with future compliance than the filing of amended returns. *This primarily is due to the timing of the subsequent year extract. We know from information copies of amended returns mailed to NFL- DORA that not quite all amended returns had been processed as of the master file extract date.*

**Chart 7. Percent of Amended Returns Filed by Group**



**Conclusions and Actions Taken**

Conclusions

*Composition of Population*

The research results demonstrate the roughly 3.22 million duplicate SSN population is composed largely of two groups. One group (1.17 million) consists of two parents filing separately, each claiming one or more of their children as a dependent and/or qualifying child. The other (1.49 million) consists of one parent (or parents in the case of a joint return) claiming a dependent child and that dependent filing a return that claims a personal exemption for him-/herself. With minor exception, data do not exist in IRS databases to determine who is entitled to the SSN claim and who is not.

This population also turns over rapidly. Our research indicates approximately two-thirds of the TY 1995 population was not in the TY 1994 population. This limits the effectiveness of any treatment involving one-to-one contact. Such contact with the duplicate SSN users would impact only the remaining one-third of the identified population that would repeat a duplicate

SSN claim. In addition, since so many taxpayers enter the population every year, absent development of a preventative treatment (such as a tax law, tax form, or schedule change), a yearly application of a treatment would be necessary to reduce the size of the duplicate SSN population from year to year.

#### Effectiveness of Notices and Examinations

Despite this inherent limitation (arising from the low repeater rate), the soft notices were nevertheless effective at improving compliance. For example, the notices reduced the instances of subsequent duplicate SSN claims by 15.7 percentage points for taxpayers overall, compared to the control group. Also, there was an indication the notices were slightly more effective on this dimension than correspondence examination. In addition, educational notices are less threatening and burdensome to taxpayers than examinations and such examinations cost more than treatment by notices.

Nevertheless, Examination contact letters also improved compliance by reducing subsequent duplicate SSN claims among the repeater sub-population tested. In addition, not only do the examinations result in improved future compliance (prevention), they also result in recovery of lost revenue. The revenue potential likely is greater in examinations because the number of taxpayers disclosing deficiencies on previously filed returns is greater than the apparent small number of voluntarily amended returns likely to arise from the notice treatment.

#### Operational Actions Taken

The test results described above, as well as earlier research findings in the duplicate SSN area, gave rise to a number of promising program adaptations (initiatives). These initiatives were targeted to specific segments of the duplicate SSN population and incorporated into the National Strategies within the Operations Plan -- the comprehensive, multi-year planning document designed to improve compliance through innovative ideas for the IRS functions under the Chief Operations Officer. Those initiatives

included the "Duplicate TIN Repeater Project," which involved correspondence examinations for some 145,000 TY 1995-1996 duplicate SSN repeaters claiming the earned income tax credit, and a nationwide soft notice mailing to another 2.3 million taxpayers claiming a duplicate SSN for TY 1996. North Florida DORA staff currently are measuring the effectiveness of those operational program initiatives.

---

#### **About the Author(s):**

**Ivette Alamo-Tirado** began her career as an operations research analyst with the IRS in 1991, and is presently an ORA with the North Florida District Office of Research and Analysis. She holds a B.S. degree in industrial engineering from the University of Puerto Rico, an M.S. degree in engineering administration and systems analysis from George Washington University, and an M.B.A. degree from Turabo University, Puerto Rico.

**Robert Holmes** is a program analyst Team Leader in the North Florida District Office of Research and Analysis. He began his career with the Internal Revenue Service in 1971 upon graduating from Jacksonville University with a B.S. degree in accounting. He is also a certified public accountant.

# High-Range Corporation Return Workload Selection System Development

By James A. Wilhelm

*Prior to this research, IRS's Examination function did not have a mathematical model to help it prioritize its high-range corporate workload. Examination defines high-range workload as audit work associated with corporations and partnerships with over \$10 million in total assets not designated as part of the Coordinated Exam Program (CEP) (i.e., its large case program). This article discusses the development of an operational selection model for the high-range corporate income tax return (Form 1120) workload. When the Discriminant Analysis System (DAS) scoring model becomes fully functional for these types of returns, Examination has the potential to assess almost \$1.2 billion in additional tax revenue per fiscal year. Additionally, the DAS model will help Examination reduce taxpayer burden by reducing by half the number of such corporations audited. The ability of this scoring system to identify more productive cases for examination and reduce taxpayer burden led to its inclusion as one of IRS's ten National Strategies designed to increase compliance through innovative approaches. With initial implementation in fiscal year 1998, Examination used the DAS scoring model to identify 25 percent of its high-range Form 1120 workload.*

## Background

After a 1995 General Accounting Office (GAO) audit, IRS Examination officials stated they "want better systems for selecting and classifying returns."<sup>15</sup> They were referring specifically to the high-range corporate and partnership returns. These high-range returns include those for corporations and partnerships with \$10 million and over total assets not designated as part of the large case program, or CEP. Approximately 53,000 such returns are filed each year.

Currently, Examination does not have a mathematical selection system for the high-range returns. As a result, selection of returns for examination in this category is not very effective. For fiscal years (FYs) 1993, 1994, and 1995, 47.7 percent of all closed high-range examinations resulted in no additional assessments from the audits (i.e., no tax liability increase over what the taxpayers reported on their returns). This suggests current audit selection is not much better than random selection. Examination needs more effective selection systems for the high-range returns to ensure it examines the most non-compliant returns and improves yield in this area.

## Objective

The objective of this project was to develop a national scoring model for all non-CEP high-range corporate returns. This model would prioritize returns for the IRS Examination function based on their probability of being profitable to audit (as defined later in this article). That is, the profitable to audit (PTA) cases will have a high rank and the non-PTA cases will have a low rank.

For this effort, we decided to look at various alternative statistical modeling techniques, including discriminant function analysis and regression. With any of these techniques, our goal was to predict an outcome (i.e., PTA or not PTA) based on various predictor variables -- i.e., tax return line items. Therefore, each of these techniques required a database of returns that contained both prior audit results and tax return data. The model discussed in this report was developed to operate within the current IRS return submission, processing, and examination systems.

## Data Used

The data we used for model development came from the following three sources.

- 1) Midwest Automated Compliance System (MACS) MACS contains transcribed tax return data for all high-range

---

<sup>15</sup> Tax Administration: Audit Trends and Taxes Assessed on Large Corporations (GAO/GGD-96-6, October, 1995, p. 9)

corporations. We used processing year (PY) 1992, 1993, and 1994 line item tax return data. (“Processing year” refers to the year in which the returns were filed.)

- 2) Office of Foreign Business Study (OFBS) This database is produced by the Statistics of Income (SOI) Division at the IRS. It contains line item information on corporation returns with over \$10 million in total assets<sup>16</sup>. We used PY 1989, 1990, 1991, 1992, and 1993 data to create a long term portrait of the businesses under study.
- 3) Examination Closed-Case Database This database contains various data for all completed examinations for a given fiscal year. A component of the database is audit-related information for these returns. We used FY 1993, 1994, and 1995 closed-case data.

We used these three databases to create our development database. We first merged the MACS and SOI data with the Closed-Case data in order to match tax return line item data with their corresponding Examination results. We then eliminated all returns without corresponding Examination results. We also eliminated all returns designated as CEP or duplicate returns (i.e., returns filed for the same tax period). We had 16,415 returns remaining in our database after this merging and elimination.

The 16,415 returns represented 51.5 percent of the 31,855 high-range returns examined and completed by Examination in FYs 1993, 1994, and 1995. They were filed on or after PY 1989. The remaining 15,440 returns were filed before PY 1989 (e.g., some returns were filed in 1968). We could not match any of these returns since

<sup>16</sup> The SOI database contained information on 30 percent of AC 219 returns and all AC 221 through AC 225 returns. Activity Code (AC) is Examination’s way of categorizing its workload. It is based on size of total assets. -AC 219 includes returns with assets of \$10,000,000 < \$50,000,000. AC 221 includes returns with assets of \$50,000,000 < \$100,000,000. AC 223 includes returns with assets of \$100,000,000 < \$250,000,000. AC 225 includes returns with assets of \$250,000,000 or more.

we no longer have the tax return line item data for these pre-PY 1989 returns. Since we needed line item information for this project, we could not use the information contained in these pre-PY 1989 cases. The characteristics of the pre-PY 1989 returns may or may not be different from those of the post-PY 1989 returns. For purposes of model development, we assumed the differences were minimal.

The methodology for developing workload selection systems is straightforward. A key assumption was our database consisted of a random sample of high-range corporate returns.<sup>17</sup> We then developed a selection model using both a logistic regression approach and a discriminant analysis approach. The following discussion provides more details.

#### Determining “PTA” Criteria

Before we started model development, we needed to determine the criteria for a “profitable to audit”<sup>18</sup> return. A definition was essential for each of the statistical techniques we used for model development. In the high-range area, Examination has defined the PTA criteria as “\$2,000 of audit results per hour” (i.e., a \$2,000 increase to a return’s tax liability resulting from every hour expended on an audit). Using this definition, slightly over 10 percent of the returns (1,722 returns) in our project database were PTA returns.

#### Development Versus Test Files

Once we determined the PTA criteria, we split our database into a formula development file and a test file. This “splitting” was done so we could evaluate, or measure, the predicted effectiveness of the developed models on an

<sup>17</sup> Ideally, the examination results we use to develop selection systems should be from a random sample of the population. This is so we can feel confident that the results are statistically representative of the population allowing us to develop an unbiased selection system. In reality, we do not have results from a random sample. However, Examination believes that current returns selection in this area is not much better, if any, than random selection. Therefore, we will make the assumption that exam results will reflect what we would see if Examination audited a random sample of these cases.

<sup>18</sup> For our “profitable to audit” definition, we were concerned only about the change in tax (audit results). We did not consider the change to the Net Operating Loss Deduction (NOLD) as a potential PTA definition. However, other IRS research on “Revenue Protection” uses the NOLD as a potential PTA definition.

independent test file of cases not used in model development. To further enhance this evaluation, we required our two split files to satisfy three conditions to ensure they were as similar as possible. The conditions were:

- 1) the distribution of the development file returns be similar to the distribution of the test file returns (e.g., the percent of returns in the development file with total assets between \$50 million and \$100 million had to be approximately equal to the percent in the evaluation file);
- 2) the average tax change of the returns in the development file had to be similar to the average tax change of the returns in the test file; and
- 3) the percent of PTA cases in the development file had to be similar to the percent of PTA cases in the test file.

With these “splitting” conditions in mind, we let SPSS (our statistical analysis software package) randomly select 70 percent of the 16,415 returns in our database for the development file. We wanted the development file to contain a large portion of the project database in order to increase the chances of identifying data relationships between examination results and tax return line items. The remaining 30 percent (4,899 examined returns) constituted the test file. For all model development, we used data from the development file. We then used the model to score and rank the returns on the test file to predict how well the model would perform if Examination used it to select returns.

#### Determining Sampling Weights

Our final data-preparation task was to determine the sample weights. Since we assumed the available data was a random sample of the Form 1120 population, we needed a way for our sample data to represent the high-range corporate population for the most current year. We noted there is a multi-year lag between when a return is filed/processed and when it is selected and examined. Data from Examination indicated the total population consisted of 53,394 Form 1120 filers in PY 1994, the most

current processing year for which there are completed corresponding examinations. We wanted our “samples” to represent this processing year. Hence, the weight we assigned to each return in our overall file was 3.25 (i.e., 53,394 divided by 16,415). The weight we assigned to each return in our test file was 10.899 (i.e., 53,394/4,899).

#### Model Development

After we completed this data preparation, we developed the first set of models using only tax return line item data relative to dollars. For the first set of models, we found the discriminant function models performed better than the logistic regression models. We found the logistic models correctly classified the non-PTA returns. However, the discriminant function models did a better job of classifying both the non-PTA returns and the PTA returns. Since it was more important to classify correctly the PTA returns (from an Examination standpoint), we continued our efforts using the discriminant analysis technique.

We then considered “non-dollar” return line items as we continued our formula development. Two examples of non-dollar line item variables are Accounting Code, which reflects the type of accounting method used by the taxpayer, and Principal Business Activity (PBA) Code, which reflects the taxpayer’s type of business.

Next, we developed models for sub-segments of our population. The reason for this segmentation was to produce more refined formulas. Using our same 70 percent development/30 percent test approach, we developed segmented models for Examination Activity Codes (i.e., groups by size of total assets), major PBA categories, and combinations of the two. For each sub-segment, we developed both discriminant function models and logistic models. We again found the discriminant-based models outperformed the logistic models. We also found that grouping the returns into certain distinct PBA Codes was the best grouping strategy of those considered. That is, models that distinguish specific groupings by PBA Codes did the best job of correctly classifying both the non-PTA returns and the PTA returns.

Our final step in the development process was the inclusion of financial ratios and indicators in the model development. These financial variables were either basic economic financial ratios or financial indicators used by the IRS. We developed discriminant-based models, for PBA Code groupings, using all transcribed line item data (both dollar and non-dollar) and the generated financial data. These models outperformed all other models (e.g., those without groupings by PBA Code) in classifying the PTA returns as PTA returns. They also did a better job of classifying the non-PTA returns as non-PTA returns.

### Evaluation Approach

Our model development resulted in the creation of two alternatives for scoring and ranking returns for examination. The first alternative (Georgia DORA Selection I) was a one-formula model. That is, one discriminant function formula was applied to the entire high-range corporation return population. The second alternative (Georgia DORA Selection II) was a two-formula model. This model consisted of two discriminant function formulas, Formula A and Formula B, where Formula A applied to returns with one particular set of major PBA Codes and Formula B applied to all other returns. For Formulas A and B, we considered transcribed data, financial ratios, and financial indicators in the model development process.

For our evaluation, we replicated as closely as possible a typical fiscal year of examinations. Our definition of a typical year of examinations was the expenditure of 1,382,160 "Examination hours." We use 1,382,160 hours as a representation of a typical year for two reasons. First, it represents the average number of hours Examination spent auditing (and ultimately closing) high-range Forms 1120 tax returns in FY 1993, 1994 and 1995. Second, an hour representation tempers "complexity of the audit" dilemmas, "case grade" dilemmas, and "indirect staff time" dilemmas caused by a standard "case" fiscal year representation.

We used returns from the 30 percent test file for our evaluation<sup>19</sup>. These returns were **not** used for formula development. The sole purpose of these "test" returns was to determine how well our formulas identify a "profitable to audit" return. This type of evaluation is commonplace for statistical model development.

Our evaluation used four different approaches for selecting returns to examine.

- 1) Current Examination Selection  
For this approach, we averaged the results for three fiscal years (1993, 1994, and 1995) of examinations. These averages will be the best estimate of how Examination currently conducts its high-range corporate program.
- 2) Georgia DORA Selection I  
In this approach, we applied our best single-formula model to all 4,899 returns on the test file. We sorted and ranked the returns in descending score order. We then selected returns from our ranked listing until we accumulated, at most, 1,382,160 weighted hours. When we expended these Examination hours, we recorded the predicted audit results. We believe these results represent what we could expect in a typical fiscal year in Examination, if Examination used the one-formula model.
- 3) Georgia DORA Selection II  
For this approach, we applied both Formulas A and B to their respective "test file" returns. We sorted and ranked both sets of returns in descending score order. For this set of formulas, we needed a relative predictive measure for combining the two scored groups into one sequenced listing. To help determine this sequence, we

<sup>19</sup> To be as conservative as possible, we eliminated one "outlier" case from the test file. This case took 2,560 hours and yielded over \$105 million in tax change.

decided to use a probability SPSS assigns to each scored case. This probability represents the likelihood of the return being profitable to audit. The higher the probability, the more likely the return is PTA. Thus, we created a single ranking of all the cases from the highest to lowest under the Selection II approach using the SPSS probability assigned to each return.

We then selected the appropriate returns from our ranked listing under Selection II until the total hours we accumulated equaled 1,382,160 weighted hours. When we expended these examination hours, we recorded the results. These results are our best representation of a typical fiscal year in Examination, if Examination used the two formulas to select returns.

4)

Perfect Selection

This approach represents the "ideal" scenario. To determine perfect selection, we sorted and ranked the 4,899 returns in the test file in descending audit result order. That is, the largest tax change (per hour) return is first, the second largest tax change return is second, etc. We then went down our ranked listing until we accumulated 1,382,160 weighted hours. When we expended these exam hours, we recorded the results.

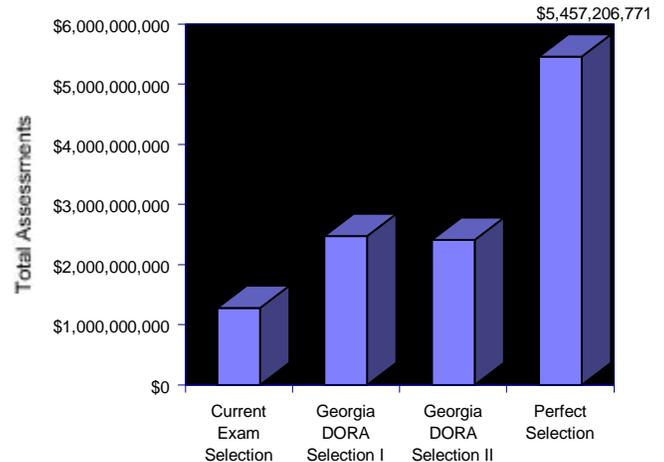
To better understand the ranking abilities of the various selection approaches, we evaluated them using five different measures.

1. Total Audit Results
2. Dollars per Hour
3. Profitable to Audit Rate
4. No Tax Increase Rate
5. Number of Cases Worked

Total Audit Assessments Can Be Increased by Nearly \$1.2 Billion

One measure of the success of the Examination function is its ability to assess accurately additional tax after audit. As seen from Figure 1, both the Georgia DORA Selection models can increase total Examination assessments. We estimate that Examination could generate an additional \$1.197 billion (i.e., \$2,474,289,051 - \$1,277,115,840) in assessments per year if the Georgia DORA models were used to select returns for examination<sup>20</sup>. Based on our matched files, this is a 94 percent improvement over the historic Examination selection procedure. However, there is still room for improvement. The Georgia DORA models only identify half the total optimal tax change under perfect selection.

**Figure 1. Comparison of the Georgia DORA Selection Models with Current Examination Selection and Perfect Selection for Total Dollars Assessed (Based on Projected/Actual Total Audit Results)**

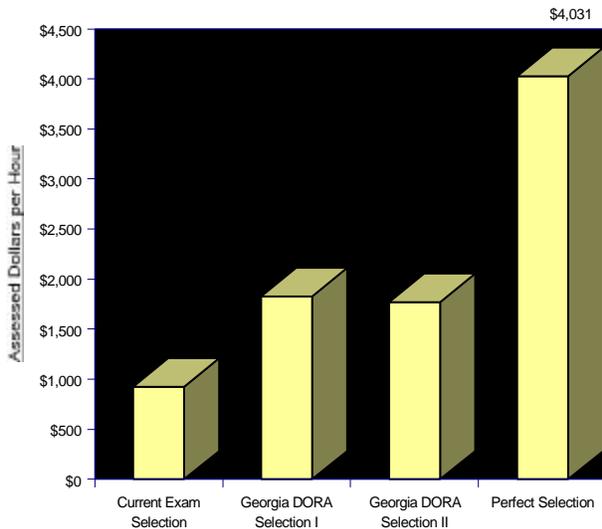


<sup>20</sup> Total improvement will **not** be realized until the fiscal year in which all discretionary returns closed were selected for Examination using the models. Thus, improved results will accumulate over time.

Dollars Per Hour Can Be Improved by 97 Percent

Another measure of interest is the amount of assessments an examiner can produce in an hour. This commonly is referred to as “dollars per hour.” The average dollar per hour rate for high-range corporations based on our matched files is \$924. By using the Georgia DORA Selection I model, Examination could realize over \$1,800 per hour (see Figure 2). This is a 97 percent improvement in dollars per hour. However, as we saw in the total audit results section, there is room for even more improvement.

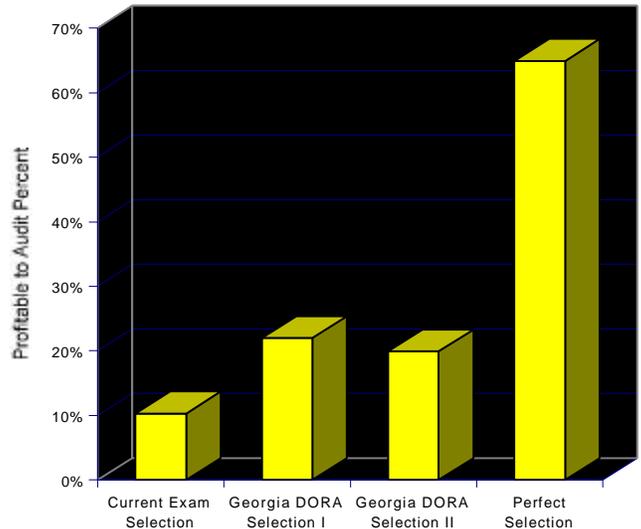
**Figure 2. Comparison of the Georgia DORA Selection Models with Current Examination Selection and Perfect Selection for Assessed Dollars Per Hour (Based on Projected/Actual Dollars Per Hour)**



Profitable To Audit Rate Can Be Improved by 100 Percent

Another goal for Examination is to work only profitable to audit returns. From Figure 3 we see the historic PTA rate for Examination is 10.31 percent. That is, almost 11 returns out of 100 audited have a tax change in excess of \$2,000 per hour. We also note the PTA rate for both of the Georgia DORA models is around 20 percent. Thus, the use of a Georgia DORA Selection model would improve the PTA rate by about 100 percent.

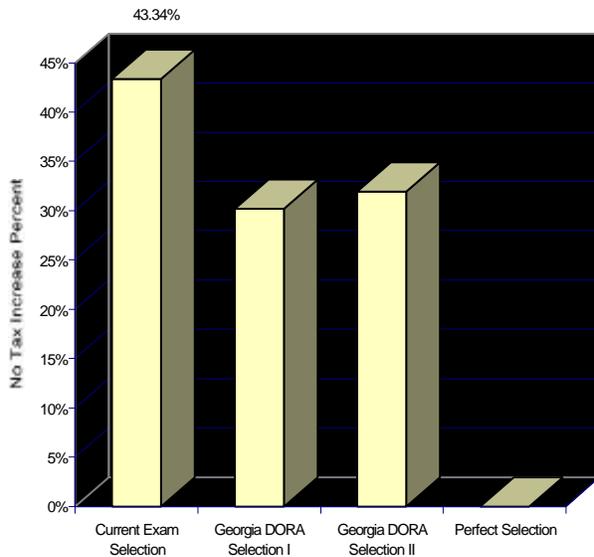
**Figure 3. Comparison of the Georgia DORA Selection Models with Current Examination Selection and Perfect Selection for PTA Percent (Based on Projected/Actual Profitable to Audit Rate (in percent))**



No-Tax-Increase Rate Can Be Decreased by 30 Percent

Another measure to Examination is the no-tax-increase (NTI) rate. Examination does not want to devote resources on audits that do not generate revenue. This wastes the taxpayer's time and wastes Examination resources. We wanted our models to select fewer NTI returns for examination. It appears the Georgia DORA models will reduce the number of these erroneously selected returns in the audit inventory. Depending on which model is used, we see from Figure 4 the Georgia DORA models can reduce the NTI rate by anywhere from 26 to 30 percent.

**Figure 4. Comparison of the Georgia DORA Selection Models with Current Examination Selection and Perfect Selection for the No Tax Increase Percent (Based on Projected/Actual No Tax Increase Rate (in percent))**



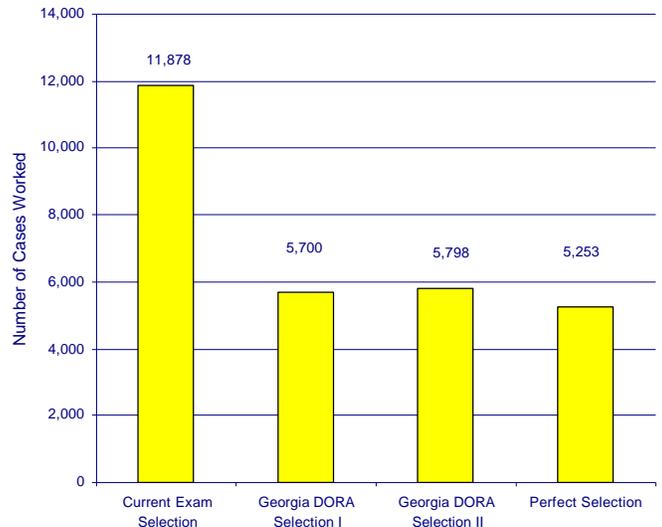
Reduced Number of Cases Worked/Taxpayer Burden

Another measure important to Examination is the allocation of cases to its staff and the associated burden on taxpayers. From Figure 5 we see Examination historically closes almost

12,000 high-range corporate audits each fiscal year. By using a Georgia DORA model, Examination could expect to audit up to 52 percent fewer returns, since better returns will be selected for audit. This reduction would impact work assignments both between and within IRS districts.

The reduced number of cases also reduces taxpayer burden. Using the Georgia DORA models, Examination should reduce overall taxpayer burden in this Form 1120 group by one-half. Also, since the Georgia DORA models will identify fewer no-tax-increase returns, Examination further could reduce the number of taxpayers experiencing an audit that results in no substantial change to their tax liability. This reduction will go from 5,148 audited no-tax-increase returns (i.e., 11,878 x .4334) to a projected 1,722 (i.e., 5,700 x .3021). Therefore, by using the Georgia DORA models, Examination could reduce taxpayer burden on this dimension by examining 70 percent fewer no-tax-increase returns.

**Figure 5. Comparison of the Georgia DORA Selection Models with Current Examination Selection and Perfect Selection for Number of Cases Worked (Based on Projected/Actual Number of Cases Worked for a Fiscal Year)**



This research showed Examination could use the Georgia DORA selection models to identify better workload in the high-range Form 1120 corporate area. The new selection models promised not only more productive use of Examination staff, in terms of assessments, but also significant reductions in taxpayer burden. As a result of this promising innovation, the high-range Form 1120 scoring systems were adopted as one of the National Strategies. The National Strategies are a key component of the Operations Plan, the core multi-year planning document for the IRS's Chief Operations Officer.

In FY 1998, Examination began initial implementation of the high-range Form 1120 Discriminant Analysis System (DAS) model in a limited test environment. Examination examined approximately 1,200 high DAS scored cases. This represents 25 percent of Examination's FY 1998 high-range Form 1120 workload. Most of these large corporate examinations take several years to close. Therefore, the results of the test will not be available for some time.

---

**About the Author(s):**

**James A. Wilhelm** is an operations research analyst in the Georgia District Office of Research and Analysis. He received his M. S. in statistics from West Virginia University. He has been with the IRS since 1988.

# Predicting Estate Tax Filings and Taxable Gifts

By Jonathan Feinstein and  
Chih-Chin Ho

*We develop a strategy for predicting filing for estate tax returns by integrating individual wealth and mortality statistics into a model framework, based on two non-IRS survey datasets: the Health and Retirement Study (HRS) and the Assets and Health Dynamics Among the Oldest Old (AHEAD). Both datasets provide extensive economic, financial, demographic, and health information about older Americans.*

*We use the HRS and AHEAD datasets for estimation of household assets subject to estate taxation and develop a prediction of estate tax filings among decedents who exceeded 50 years of age in 1992. A comparison of our predictions with the IRS's Statistics of Income (SOI) Division estate tax filing statistics provides an estimate of estate tax nonfilers. Our findings suggest very few nonfilings of estate tax returns for decedents age 80 and above, for which the bulk of estate taxes are paid.*

*Since the lifetime giving of gifts is considered a preferred tool for estate planning among wealthy individuals, we also develop a strategy for estimating the incidence and magnitude of noncompliance with the gift tax. We tabulate responses to questions about gifts in HRS and AHEAD and construct preliminary estimates of taxable gifts for 1992. A comparison of our results with IRS gift tax reporting statistics provides an estimate of unreported gift giving above the threshold for paying tax. Our findings indicate more substantial underreporting of gift tax liability. However, our estimation methodologies for both the estate and gift tax areas reflect new approaches with some known weaknesses. As a result, the tax gap estimate in this article should be viewed as tentative, with future research planned.*

## Data Sources

Both the Health and Retirement Study (HRS) and the Asset and Health Dynamics Among the Oldest Old (AHEAD) are panel datasets funded by the National Institute of Aging of the Department of Health and Human Services and administered by the Institute of Social Research of the University of Michigan. They provide extensive economic, financial, demographic, and health information about older-age population of the United States.

The HRS began in 1992. Beginning in that year, and every second year for 12 years thereafter, each household in the HRS has been interviewed and the household's respondents have been asked numerous questions about household income and assets, employment and retirement status of individual household members. To be selected for participation, the head of a household must have been between 51 and 61 as of 1992.

AHEAD is a longitudinal study of U.S. population cohorts born prior to January 1, 1924. From October 1993 to April 1994, AHEAD Wave 1 interviews were conducted with national

samples of these age 70-plus individuals and their spouses about major transitions in their health and financial situations. The longitudinal study plan specifies a full-scale re-interview of the AHEAD panel every second year beginning in 1995.

We extracted information on age, race, sex, marital status, life insurance, and assets from the HRS and AHEAD datasets. We then performed numerous data manipulations in order to impute missing data.

There is considerable overlap between the information recorded in HRS and that in AHEAD. Both datasets contain information about household composition, household income and assets, including pension assets, employment and retirement status of individual household members, health care use and costs, health insurance and life insurance of individual household members, housing, and the household's economic relationship with other non-household family members.

However, despite the many similarities, there are some differences between the datasets in the information they gather. One important difference for work on estate

taxation is that AHEAD asks about trust assets, while HRS does not. We discovered that trust assets are a significant proportion of total assets for many households in the top 10% of the wealth distribution.

We computed several different measures of total household assets, differing in how we interpreted the trust asset variables. We report our results for the case that is biased in favor of the largest amount of trust assets, the case that we believe is associated with the most plausible interpretation of the respondents' answers to the series of questions they were asked about trust assets.

An estate tax return, Form 706, must be filed with the IRS for any decedent whose estate satisfies certain conditions. For tax year 1992, the most important condition is that the gross value of the estate exceed \$600,00; however, starting in 1998, recent tax legislation increases the threshold in steps to \$1 million by 2006.

For all decedents meeting the filing requirements, an estate tax return must be filed within 9 months of the date of death, although an extension of six months is allowed. Although a return must be filed for every decedent whose estate value exceeds this threshold, tax is not owed on all such estates. Certain deductions may be subtracted from the gross value, and tax is owed only on the net value of the estate.

Further, if a decedent is married an unlimited marital bequest may be made to his (her) spouse, in which case tax is owed only on the net value of the estate not included in the bequest in excess of \$600,000. If a decedent is widowed and his or her spouse made a marital bequest, tax is owed only on that portion of the net value in excess of \$600 thousand.

The HRS and AHEAD datasets can be used to predict estate filings among decedents who exceeded 50 years of age in 1992. Our strategy is to divide the post-50 population into mutually exclusive cells, based on age, sex, and marital status groups.

The five age brackets include one group taken from HRS (50-61), three groups from AHEAD (70-74, 75-80, 80 and over), and the 62-69 group interpolated between HRS and AHEAD. The two sex groups are male and female; and

the two race groups are white and non-white. The four marital status groups are married, widowed, divorced, and single and all others. In total, there are 80 separate cells.

We construct a measure of household assets equal to the HRS/AHEAD household net worth variable. Then we construct individual estate asset values. For non-married individuals, this equals household net worth plus life insurance; for married heads of household and their spouses, it equals one-half of household net worth plus relevant life insurance.

We compute the predicted number of estate tax filings in each cell in the following three steps:

1. We take the number of death in 1992 directly from mortality statistics provided to us by staff at the National Center for Health Statistics (NCHS).
2. We calculate the fraction of individuals for which the estate value exceeds the 1992 estate filing threshold of \$600,000, using the estate asset value variable based on information from HRS or AHEAD as described earlier. This fraction is weighted and therefore refers to the U.S. population in this cell.
3. We multiply the fraction from Step 2 by the number of deaths determined in Step 1, thus providing an estimate of the number of deaths for which it is predicted that an estate tax return should be filed.

Table 1 lists three estimates for each cell: (1) the number of deaths in 1992 reported by NCHS, (2) the proportion of individuals for whom estate value exceeds the filing threshold, and (3) the number of estate filings expected for 1992. Note that we compute predicted estate tax filings, not filings for which estate tax is due or for net tax liability. In particular, for decedents who were married, estate tax returns must be filed on their behalf although they need not pay any tax if they bequeath all or most of their estate to their spouse.

**Table 1**  
**Deaths, Proportions of Individuals Exceeding the Estate Tax Filing Threshold,**  
**Predicted Estate Tax Filings**  
**For Decedents Who Exceeded 50 Years of Age in 1992**

	Male White			Female White			Male Nonwhite			Female Nonwhite		
<i>Ages 51-61</i>												
Married	71,078	.077	5,473	39,159	.054	2,115	13,608	.025	340	7,582	.019	144
Widowed	4,146	.059	245	10,205	.043	439	1,662	.000	0	4,158	.000	0
Divorced	20,390	.077	1,570	11,587	.026	301	4,826	.006	29	4,744	.000	0
Never Married	11,358	.079	897	4,292	.037	159	4,744	.000	0	2,206	.000	0
<i>Ages 62-69</i>												
Married	111,213	.068	7,596	53,000	.054	2,862	15,070	.012	188	7,361	.010	74
Widowed	13,504	.071	959	32,713	.047	1,538	3,477	.010	35	8,438	.006	51
Divorced	18,975	.077	1,461	12,530	.046	576	3,753	.000	0	2,704	.006	16
Never Married	12,480	.076	948	5,810	.046	267	3,473	.000	0	1,734	.000	0
<i>Ages 70-74</i>												
Married	94,941	.059	5,602	40,288	.052	2,095	10,246	.000	0	4,150	.000	0
Widowed	18,745	.109	2,043	47,454	.059	2,800	3,620	.037	134	8,494	.014	119
Divorced	10,180	.100	1,018	8,507	.075	638	1,936	.000	0	1,380	.000	0
Never Married	8,407	.051	429	5,595	.073	408	1,862	.000	0	1,040	.000	0
<i>Ages 75-79</i>												
Married	97,514	.060	5,851	36,789	.055	2,023	8,836	.000	0	3,105	.000	0
Widowed	28,521	.055	1,569	75,266	.042	3,161	4,355	.000	0	10,585	.011	116
Divorced	7,869	.054	425	8,084	.058	469	1,342	.000	0	1,045	.000	0
Never Married	8,518	.093	792	7,677	.038	292	1,302	.000	0	928	.000	0
<i>Ages 80+</i>												
Married	148,141	.040	5,926	46,017	.032	1,473	12,137	.000	0	3,217	.000	0
Widowed	103,211	.076	7,844	356,589	.032	11,411	11,400	.000	0	33,125	.000	0
Divorced	9,033	.041	370	16,528	.018	298	1,522	.000	0	1,448	.000	0
Never Married	15,518	.095	1,474	33,454	.076	2,543	1,914	.000	0	2,308	.000	0

- In each cell, the first number is the number of deaths reported by NCHS, the second number is the fraction of individuals in that cell category for whom estate value exceeds the filing threshold, based on tabulations made using the HRS and AHEAD datasets, and the third number is the predicted number of estate tax filings.
- For the interpolated cells (Ages 62-69), deaths from the NHCS are imputed as:  
0.65 X deaths among ages 60-64 + all deaths among ages 65-69;  
Wealth fractions are computed as: .50 X fraction of individuals' estates valued above \$600,000 among ages 51-61  
+.025 X fraction of individuals above \$600,000 among ages 70-74  
+.025 X fraction of individuals above \$600,000 among ages 75-79

### Comparing Actual and Predicted Estate Tax Filings

We have combined the individual asset distribution derived from HRS and AHEAD with the mortality data provided by the NCHS to develop preliminary estimates of the number of estate tax filings expected for year of death 1992. We also compared the expected filings with the actual filings tabulated by SOI for 1992. Such a comparison generates a prediction of the number of estate tax nonfilers in 1992.

This involves comparing our estimates of predicted filing for 1992 decedents with SOI statistics about actual estate filings for 1992 decedents. Since the SOI estate filing statistics are available only for 3 age groups - 2 sex brackets - 3 marital status categories (18 cells), we collapse over relevant cells to form an 18-cell comparison.

Table 2 summarizes the analysis and provides preliminary estimates of the number of estate tax nonfilers, together with additional information gleaned from Eller (1995)<sup>21</sup>. The table divides the data into three age groups (51-69, 70-79, 80 and over), two sexes (male, female), and 3 marital status groups (married, widowed, others).

Table 2 lists five statistics for each cell: (1) the predicted number of estate filers, (2) the actual number of estate filers reported by SOI, (3) the predicted number of estate nonfilers, (4) the total gross estate value reported by the filers, and (5) the net estate tax reported by the filers.

The number of predicted estate tax filings is significantly greater than the actual number for decedents below the age of 80, but is approximately equal to the actual number for decedents 80 and above. If the finding is correct it suggests that filing noncompliance is concentrated among families of younger decedents. However, since the AHEAD data is thin for individuals 80 and above, we may be under-predicting estate values for this age group.

Approximately two-thirds of estate tax is paid on estates of decedents who were 80 or above.

<sup>21</sup> Eller, Martha, "Federal Taxation of Wealth Transfers, 1992-1995," *Statistics of Income Bulletin* (Winter 1997), pages 8-63.

Approximately one-third is paid for a single category, widowed females 80 and above. It is clearly important to learn more about this group.

Overall, the preliminary estimates suggest that there may not be much of a nonfiler problem for estate taxes. The estimates do predict nearly twice as many filings for decedents below the age of 80, but much of it may be due to the failure to take into account the relationship between wealth and mortality. The estimates suggest that there are very few nonfilings for decedents age 80 and above, for which the bulk of estate taxes are paid.

### Estimating Taxable Gifts

A household must file a gift tax return, Form 709, with the IRS and pay the gift tax if during the year it makes a gift or gifts to an individual with total monetary value exceeding the household's gift tax threshold. In general, the donor of the gifts must file Form 709 on or after January 1 but not later than April 15 of the year following the calendar year when the gifts were made. Tax is due on the excess of the amount above the threshold. A separate line item must be recorded for each individual donee for whom total monetary value of the gift or gifts exceeds the threshold. The tax threshold is \$20,000 for a married couple and \$10,000 for divorced, widowed, and never married individuals.

Both HRS and AHEAD ask about gifts in reasonable detail. HRS asks the household respondent about gifts given to parents and children during the past year, allowing for information about the value of gifts given to up to four different children. For each child to whom a gift was given, HRS asks whether the gifts were in part for education or housing. AHEAD asks about gifts given to children and other individuals during the past year.

Table 3 presents statistics about gift giving as reported in HRS and AHEAD, both unweighted and weighted to reflect the U.S. population of households. The percentage of households reporting making gifts to at least one individual for which the total value of the gifts exceeds the household's tax threshold is approximately one half percent in HRS and one and one-half percent in AHEAD.

**Table 2**  
**Actual and Predicted Estate Tax Filings**  
**Reported Gross Estate Values and Net Estate Taxes**  
**For Decedents Who Exceeded 50 Years of Age in 1992**

	Age 69 -		Age 70-79		Age 80 +	
	Male	Female	Male	Female	Male	Female
<b>Married</b>						
Predicted Number of Filers	13,597	5,195	11,453	4,118	5,926	1,473
Actual Number of Filers	8,217	2,417	7,072	2,130	7,046	1,700
Predicted Number of Nonfilers	5,380	2,778	4,381	1,988	(1,120)	(227)
Reported Gross Estate Value	\$15,486 M	\$3,866 M	\$14,333 M	\$3,198 M	\$13,827 M	\$2,675 M
Reported Net Estate Tax	256 M	115 M	374 M	155 M	768 M	198 M
<b>Widowed</b>						
Predicted Number of Filers	1,239	2,028	3,746	6,196	7,844	11,411
Actual Number of Filers	513	1,155	1,538	3,084	4,977	11,825
Predicted Number of Nonfilers	726	873	2,208	3,112	2,867	(414)
Reported Gross Estate Value	\$760 M	\$1,749 M	\$2,374 M	\$4,522 M	\$9,247 M	\$19,072 M
Reported Net Estate Tax	109 M	314 M	443 M	805 M	1,727 M	3,574 M
<b>Single, Divorced, Other</b>						
Predicted Number of Filers	4,905	1,359	2,664	1,807	1,844	2,841
Actual Number of Filers	2,078	795	1,021	782	1,504	2,228
Predicted Number of Nonfilers	2,827	564	1,643	1,025	340	613
Reported Gross Estate Value	\$3,267 M	\$1,189 M	\$2,391 M	\$1,014 M	\$2,455 M	\$2,730 M
Reported Net Estate Tax	393 M	185 M	323 M	135 M	333 M	297 M
<b>Total</b>						
Predicted Number of Filers	19,741	8,582	17,863	12,121	15,614	15,725
Actual Number of Filers	10,808	4,367	9,631	5,996	13,257	15,753
Predicted Number of Nonfilers	8,933	4,215	8,232	6,126	2,357	(27)
Reported Gross Estate Value	\$19,513 M	\$6,804 M	\$19,098 M	\$8,734 M	\$25,529 M	\$24,477 M
Reported Net Estate Tax	760 M	615 M	1,140 M	1,095 M	2,828 M	4,069 M
<b>Total</b>	<b>Male</b>	<b>Female</b>	<b>Age 69 -</b>	<b>Age 70-79</b>	<b>Age 80 +</b>	<b>All</b>
Predicted Number of Filers	53,218	36,428	28,323	29,984	31,339	89,646
Actual Number of Filers	33,696	26,116	15,175	15,627	29,010	59,812
Predicted Number of Nonfilers	19,522	10,312	13,148	14,356	2,330	29,834
Reported Gross Estate Value	\$64,140 M	\$40,015 M	\$26,317 M	\$27,832 M	\$50,006 M	\$104,155 M
Reported Net Estate Tax	4,728 M	5,779 M	1,375 M	2,235 M	6,897 M	10,507 M

We have used the HRS and AHEAD statistics to develop a very rough estimate of aggregate gift reporting and tax liability for households in which the head is older than 50 years of age, for tax year 1992. The combined figures from HRS and AHEAD suggest that at least 329,000 households should have reported gifts and paid the gift tax in 1993, with total value of gifts on which tax is owed equal to \$7.2 billion. With an average marginal tax rate of 31%, the tax liability on these gifts equals \$2.23 billion.

The Internal Revenue Service Data Book 1993-94 reports that 211,000 gift tax returns were filed in fiscal year 1993 with the aggregate tax paid equal to \$1.46 billion. A comparison with our results suggests about one third of all large gifts (above the threshold for reporting and paying tax) are not reported, and the associated gift tax liability underreporting is likely to exceed \$0.77 billion.

### **Methodological Qualifiers and Future Refinements**

The most significant methodological weakness of the estimation procedure used to generate the preliminary results of estate tax nonfilings is the failure to take into account the relationship between wealth and mortality. A large body of literature suggests that individuals of higher socioeconomic status (SES) live longer than individuals of lower status. Most of these studies use current household income or educational status as proxies for SES, variables that are not directly relevant for estate tax analysis. Recently Attanasio and Hoynes (1996)<sup>22</sup> have investigated the relationship between wealth and mortality using data from Survey of Income and Program Participation (SIPP). Their findings confirm that individuals residing in wealthier households have a significantly lower mortality rate than individuals residing in poorer households.

Another area of extension relates to the impact of marital status on mortality, since married persons typically leave all or most of their assets to their spouse, and therefore owe little if any

<sup>22</sup> Attanasio, Orazio and Hilary Hoynes, "Differential Mortality and Wealth Accumulation," *National Bureau of Economic Research Working Paper No. 5126* (May, 1995).

estate tax. Lillard and Waite (1995)<sup>23</sup> indicate that marital status exerts an important impact on mortality. From this perspective, incorporating the effect of marital status on mortality can improve estimates of estate tax.

In future research efforts, we plan to use data from waves 2 of HRS and especially AHEAD to estimate models of relationship between wealth, marital status, and mortality. The estimates of such models can be used to develop more precise predictions about the proportion of deaths among individuals in each age-sex-race-marital status cell for which the corresponding estate value exceeds the filing threshold. In particular, the model estimates can be used to generate wealth dependent mortality curves, thereby allowing a correction to be made for the fact that wealthier individuals face a lower mortality rate.

Since the lifetime giving of gifts is considered a preferred tool for estate planning among wealthy individuals, more careful analysis of responses to gift giving related questions in HRS and AHEAD can shed insight on the relationship among household wealth, individual health condition, and gift giving.

As a result, we plan to explore gift statistics from waves 2 in HRS and AHEAD in a life cycle framework to analyze various tax planning motives behind gift giving, such as estate planning or income transfer, after controlling for household wealth, family structure, and individual health and demographic factors.

### **Conclusions**

By combining data from HRS and AHEAD with mortality statistics provided by the National Center for Health Statistics, we have constructed preliminary estimates of the number of estate tax filings expected for 1992, by age, marital status, race, and sex. A comparison of our estimates with SOI statistics on estate tax filings for 1992 has enabled us to generate preliminary estimates of the number of estate nonfilers. Our conclusion is that there are not many estates for which an estate tax return should have been filed for 1992, but was not

<sup>23</sup> Lillard, Lee and Linda Waite, "Till Death Do Us Part: Marital Disruption and Mortality," *American Journal of Sociology* (March, 1995), pages 1131-1156.

**Table 3**  
**Related Gift Statistics Based on the HRS and AHEAD Datasets**  
**For Heads of Household Who Exceeded 50 Years of Age in 1992**

	Number of Households Making at Least One Gift		Total Number of Gifts	Total Value of Gifts		Total Value of Gifts on which Tax is Owed
	Unweighted	Weighted		Unweighted	Weighted	
<b>HRS Dataset</b>						
Gifts Greater than Tax Threshold	29 (.38%)	83,000 (.47%)	34	\$3.0 B	\$1.6 B	
In Part for Education	-	-	24	\$2.3 B	-	-
In Part for Housing	-	-	7	1.0 B	-	-
Gifts Equal to Tax Threshold	19 (.25%)	58,000 (.33%)	23	\$1.0 B	0	
In Part for Education	-	-	11	\$0.5 B	-	-
In Part for Housing	-	-	4	0.2 B	-	-
All Gifts	1,364 (18%)	3,300,000 (19%)	2,050	\$16.2 B	\$1.6 B	
<b>Ahead Dataset</b>						
Gifts Greater than Tax Threshold	90 (1.4%)	246,000 (1.4%)	99	\$9.1 B	\$5.6 B	
Gifts Equal to Tax Threshold	59 (.91%)	169,000 (.96%)	71	\$2.4 B	0	
All Gifts	1,549 (24%)	4,400,000 (25%)	1,965	\$23.5 B	\$5.6 B	

filed. As a result, the nonfiler tax gap is likely to be small.

Based on the information from HRS and AHEAD about gifts made to family members and others, we have conducted a preliminary analysis on noncompliance with the gift tax. Our findings suggest that there may be a substantial underreporting of taxable gifts. However, these tax gap estimates for estate and gift taxes are just the preliminary results from brand new estimation methodologies. The methodologies have some identified weaknesses, for which future research will be pursued in the hopes of improving the reliability of the estimates.

**About the Author(s):**

**Jonathan Feinstein** is a Professor of Economics at Yale University. He received his doctorate in economics from MIT and taught

economics at Stanford before joining the Yale faculty. He is a leading expert on public finance, economics of aging, and econometrics, and has published numerous articles in the leading economics and econometrics journals. He testified before the U.S. Congress on behalf of the IRS on tax compliance measurement issues in 1997.

**Chih-Chin Ho** is a Senior Economist in the IRS's National Office of Research. He received his doctorate in economics from the University of Michigan at Ann Arbor. He worked as a Senior Research Associate for the U.S. Sentencing Commission before joining the IRS. He has published actively in tax compliance, public finance, and econometric modeling. His most recent contributions include "An Empirical Analysis of Gift Giving in Later Life," "Toward Building a Profile of an Income Tax Nonfiler," and "Predicting Individual Income Tax Reporting Accuracy: A Trichotomous Choice Model."

# Payment Dynamics of Individual Accounts Receivable and a New Look at Risk

*The decisions of when and how much to pay on balance-due assessments are studied retrospectively for a panel of individual taxpayers. Combining data from the individual's current and prior accounts receivable status, observed tax return, and credit history, a heuristic assessment of risk is developed, and classification models are proposed. The results of such models suggest that delinquent taxpayers should be scored according to risk at the time of assessment, not after they have moved through the notice stream. In doing so, it may be possible to reduce intrusiveness on low-risk cases that would otherwise pay, and identify high-risk cases before they become a financial burden.*

## Section 1--Introduction

One of the most challenging problems facing the Internal Revenue Service is its Accounts Receivable Dollar Inventory (ARDI), the amount of outstanding taxes, penalties, and interest owed by taxpayers at a given time. At the end of 1997, gross total ARDI stood at over \$228 billion, with 52 percent of that owed by individuals.

Individual ARDI continues to grow—from \$93.6 billion at the end of 1994 to \$118.8 billion in 1997. More serious, however, is that this three-year rate of growth (26.9%) significantly exceeds the rate of growth of new individual filers during that same period (5%). Equally serious is that it also exceeds the growth in nominal U.S. personal income during that time (18.7%), which is an important factor used ultimately to determine ability to pay.

Those familiar with these statistics continue to ask the familiar questions: Why is individual ARDI growing? Who are those most likely to pay? What factors contribute to those with the greatest risk of becoming a financial burden? What constitutes risk?

This paper addresses these questions for individual accounts that file a return with a balance due. It does so by developing a new but simple approach for assessing risk based on such factors as the timing of payments, return characteristics, current and prior accounts receivable status, and credit reporting data. The analysis will guide the reader through the dynamics of payment decisions, identify factors that contribute to the likelihood of risk, and describe patterns that are ultimately of value for predicting the likely disposition of cases.

Inferences from this analysis will also motivate other questions: How much of ARDI is a systemic problem related to outdated billing practices or poor toll-free telephone access rates? To what degree do repeaters and even tax law complexities contribute to the growth of delinquencies? And perhaps most importantly, what classification system is currently in place, if at all, to accurately identify the relative risk of accounts at the time of assessment? Is it possible that such a system could be used to reduce intrusiveness on low-risk cases that would otherwise pay? Identify high-risk cases before they become a financial burden (i.e., not fully paid within one year)? And allow for a more flexible development of strategies and efficient prioritization of resources?

The remainder of this paper is organized as follows. Section 2 introduces the data used for this analysis. Section 3 examines selected summary statistics for the panel of individuals under consideration. Sections 4 and 5 use this summary information to identify and analyze subgroups of relative risk. Section 6 isolates and examines several issues related to deferrals; and Section 7 offers recommendations based on the results of this study.

## Section 2--The Data

In order to comprehensively examine payment dynamics and other characteristics of delinquent individual accounts, a panel must be constructed and followed retrospectively through time. For purposes of this analysis, a panel comprising 33,263 accounts from Maryland and the District of Columbia was developed. Each individual account in this panel has a tax delinquency due exclusively to a balance-due return that has not

been fully satisfied. The date of assessment on these accounts covers the last two weeks of May 1996.

Several points about the construction of this panel are necessary. First, this analysis investigates the payment characteristics only of balance-due delinquencies, which are the dominant source of assessment for individual receivables. Those that are the result of an audit or other source-of-assessment are not included here, and a separate analysis is recommended. Second, it would be ideal to capture *N* cases at a specific point in time, say May 15, and track the payment dynamics of these cases over one or two years. However, the proportion of cases selected on a given day represents, in effect, the probability of an assessment on that day, which is almost zero, thus resulting in very few cases to study. (Recall that the probability of a given point is undefined, but for a small neighborhood around that point it is in fact well defined.) Hence, a small neighborhood of two weeks was used. Given that these cases are being tracked for as many as three years (156 weeks), it is unlikely that the results of this analysis will be sensitive to this two-week initial condition.

Finally, it should be noted that a representative national sample could have been drawn for this study. However, such a sample would have to be designed to account not only for unique differences across geographic locations but time periods as well; not knowing these unique features might lead to inefficient design and thus poor sample representation. By analogy, any other district office could have been chosen as well, as there was no *a priori* information about taxpayers from Maryland and D.C. that prejudiced their selection (except perhaps that the overall ARDI population was large relative to other districts, resulting in a larger panel). These issues, particularly the need for cross-validation, will be discussed later.

### Section 2.1--Data Sources

This analysis combines data from the Individual Master File, Accounts Receivable Database, and Credit Reporting sources. It is admittedly limited by the lack of additional data thought to be relevant for certain subgroups, some of which will be discussed in Section 4. Notwithstanding these limitations, however, it will be shown that

these data sources contain sufficient information necessary not only to identify and describe important characteristics of payment dynamics and risk, but to predict outcomes of risk as well.

Because this is the first known effort to use third-party credit data for the purposes mentioned above, a few comments will be made. First, no specific taxpayer's records are being analyzed here; what follows is largely a descriptive statistical analysis of averages, rates, and proportions. Second, while there is evidence from this analysis that credit data may in fact improve classification accuracy for certain subgroups, its applicability is generally not widespread; as a result, there is a cost/benefit question to be addressed concerning its future use. Finally, analysis of credit data during the course of this study has left the author suspicious about certain features that appear statistically irregular, although no rigorous validation was performed to test this assumption. As with the need for cross-validation mentioned above, future research in this area using such data should proceed with caution.

### **Section 3--Descriptive Statistics**

The ultimate goal of this study is to contribute to the ability of the IRS to determine accurately which individuals are more likely to pay, when and how much will be paid, and the likely disposition of accounts that do not pay. Doing so means talking about the *probability* of payment, and hence risk groups that relate certain individuals to certain outcomes based on observable data. This section motivates the construction of such groups by exploring broad characteristics of the panel and identifying simple descriptive features that may be helpful to classification.

For the panel of 33,263 new modules used in this study, the total assessed amount was \$52 million, with a median module balance of \$713. (The median—half of the sample above, half below—is a more appropriate measure of 'average' than the mean for income amounts, which tend to be highly skewed.) About 75 percent of the cases had a beginning balance of less than \$1,500; a full 21 percent had balances less than \$250.

Just over 50 percent of all modules were fully paid within one year, paying \$21.4 million (41.2 percent) of the total beginning module balance. Why is payment within one year important? Because this is the time between filing tax returns; it will be demonstrated later that individuals who incur a new module balance before paying their old one are, all other things equal, at greater risk of not paying *at all*. Therefore, this study uses one year as the relevant time period for the analysis of payments.

Table 3.1 shows that, of those fully pay at one year, 78 percent had only one module; other features thought to be relevant for illuminating differences among these two groups are also presented.

**Table 3.1 Summary statistics for full payment within one year, by payment status**

Feature	Pay	No Pay
Median Beginning Balance	\$531	\$933
Median Income	31,705	28,761
Median Age	41	40
% One Module Only	78	41.8
% Filing Form 1040	59.5	48.7
% Remittance with Return	32.6	15.2
% Schedule C or F	21.6	18.6
% Homeowners <sup>1</sup>	37.8	25.8
% With Investment Income <sup>2</sup>	45.3	23.9

<sup>1</sup> Presence of real estate taxes on Schedule A.

<sup>2</sup> Presence of interest, dividend, Schedule D, or other capital gains.

While this table does not include an exhaustive list of factors useful for discriminating those who pay from those who do not, it appears some factors may be more useful than others. For example, all other things equal, those who don't pay will have almost twice the average module balance; those who pay are almost 2 times as likely to have investment income or just one module, and are more than 2 times as likely to attach a remittance with their balance-due return.

While these results can be used to heuristically motivate the construction of classification models, they also raise additional questions. For those who do not fully pay within one year, how much, if any, was paid during that period, and what is the disposition of the account at the end of that period? Table 3.2 reveals that nearly half of all individuals not fully paid are in an installment agreement, paying on average about 35 percent of their module balance over the course of the year. However, it appears that certain cases—in particular those in bankruptcy, currently not collectible (CNC), automated collection system (ACS), and collection field function (CFF)—are not paying down much, if any, of their original balances. Are individuals in these categories at a greater risk of not paying? If so, we might want to identify factors that contribute to predicting the likely disposition of these cases and make such information available at the time of assessment. Doing so might permit a more flexible prioritization of resources for those cases if in fact they do present the greatest risk of becoming a financial burden.

**Table 3.2 Disposition at one year of cases not fully paid**

Disposition <sup>1</sup>	% Of Cases <sup>2</sup>	% Of Ending Balance <sup>3</sup>	Median Starting Balance	Median Ending Balance
Bankruptcy	3	3.9	\$1,315	\$1,190
CNC	4.2	8.2	1,579	1,808
Deferral	29.6	7.6	407	321
Installment	48.5	50.2	1,308	966
ACS	9.8	13.8	1,336	1,340
CFF	1.9	6.3	1,654	1,903

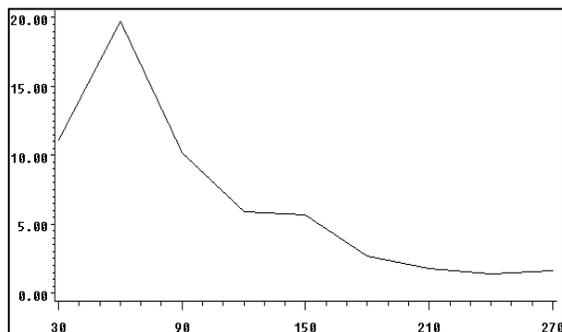
<sup>1</sup> Based on TRCAT status codes.

<sup>2</sup> Percent of only those cases not fully paid at 1 year.

<sup>3</sup> Percent of total module balance remaining.

Finally, it may be instructive to explore the *timing* of payments, and ask whether certain individuals have a higher probability of paying earlier than others; it is possible that such information may be useful in the classification sense. Figure 3.1 shows the probability of making a payment at time *T* given that one has not already been made.

Figure 3.1 Probability of making a payment versus number of days after assessment<sup>1</sup>



<sup>1</sup> The horizontal axis runs from 30 to 270 days.

This distribution reveals what might be intuitively expected: the more time that elapses before an individual makes at least one payment, the less likely it is that they will do so at *any* time after that. A crucial feature of Figure 3.1 is the rate of change of this function, which falls to approximately zero after 180 days—the typical length of the notice stream. After this time, virtually no new individuals will voluntarily begin making payments that have not already done so (unless contacted by ACS staff or a Revenue Officer).

While it is evident that the probability of making at least one payment during 180 days is a decreasing function of time from the date of assessment, this does not imply that a *response* has not been made. Many individuals with an existing delinquency, for example, respond and enter an installment agreement but make no payments on their current module because they are still paying off a previous balance. However, Table 3.3 shows that for those who do make at least one payment during 180 days, all other things equal, the probability of fully paying their balance at one year is also a decreasing function of time. This too, seems obvious: the sooner I start paying, the sooner I finish. But who is making contact and beginning payments sooner? Are there relevant data that can be used to distinguish this group from others?

As it turns out, 58.3 percent of individuals for this panel make at least one payment within 180 days after assessment. As expected, a majority (77 percent) have only one module. Also, excluding deferrals (for reasons to be discussed

below), all 21 percent of those with a balance of \$250 or less belong to this group.

The broad characteristics described thus far seem to suggest several distinct patterns: 1) individuals with very small balances may constitute little or no risk if a significant percentage fully pay within one year; 2) individuals with just one module are nearly twice as likely to fully pay than not pay, and it seems intuitively appealing to isolate such cases for further analysis; and 3) it follows that individuals with two or more modules should be separately analyzed as well.

**Table 3.3 Distribution of individuals making at least one payment within 180 days and the probability of full payment at one year**

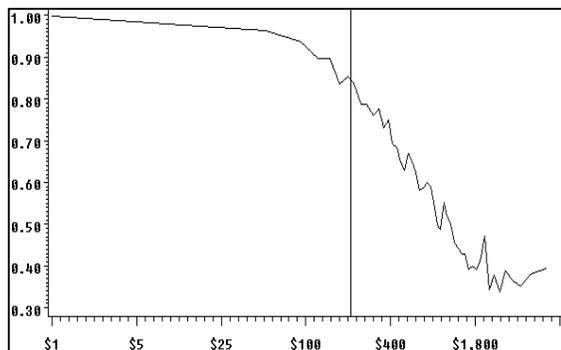
Number of Days from Assessment	% Making at least 1 Payment <sup>1</sup>	% Who Fully Pay at 1 Year	Median Starting Balance
30 Days	11.1	77.2	\$625
60 Days	19.7	75.5	848
90 Days	10.2	64	785
120 Days	5.9	64.5	699
150 Days	5.7	60	593
180 Days	2.7	49.5	694

<sup>1</sup> Number of those making at least one payment who have not yet made a payment as a percent of panel.

**Section 4--Low to Moderate Risk Groups**

Of the original panel of 33,263 cases, 6,836 (21 percent) had a beginning balance of \$250 or less. Of these, 69 percent fully paid within one year. However, of those that did not fully pay at one year, 72 percent were deferrals and 20 percent were in an installment agreement; adjusting for deferrals gives a payment rate of 88 percent for this group. This can be seen in Figure 4.1, which shows the probability of payment within one year as a function of beginning module balance. It is evident the probability of full payment is relatively constant (above 85 percent) for balances up to about \$250, after which it falls off rapidly. As a result, it appears that module balance alone is sufficient for describing the likelihood of payment for very small balances.

**Figure 4.1 Probability of full payment at one year as a function of module balance, excluding deferrals<sup>1</sup>**



<sup>1</sup> Module balance is in logarithms to accentuate the shape of the distribution; the horizontal axis shows actual dollar amounts associated with its logarithm.

Virtually all individuals with balances under \$250 either fully pay or are deferred within one year. Therefore, excluding deferrals will not affect the results. In fact, it actually helps classification efforts: whoever is not deferred will fully pay (a small percentage will be in an installment agreement). However, two questions should be asked at this point: do those in deferral status with small balances truly represent low risk? After all, one reason they are in deferral status is because they failed to respond to one or more notices, which might be considered a risk factor *per se*. And, will a significant percentage (say, 85 percent or more) in fact fully pay when the IRS eventually contacts them or visa versa? The answer to the second question will be addressed in Section 6, whereas the first question is examined in Table 4.2 below, which compares those with balances under \$250 who fully pay at one year with those who are deferred.

Although dollar amounts of \$250 or less may seem insignificant, there are features in this table that merit discussion. First, based on income alone, these two groups appear to have roughly the same *ability to pay*. If this is true, why are some responding to a notice and not others? Second, a closer look at these feature differences may lead one to conclude that the IRS may be deferring, on average, cases that are relatively more risky. For example, 19 percent of those who fully pay have just one

module versus 42 percent for deferrals. Is the IRS allowing new deferrals on top of existing ones, and if so, does this eventually increase the risk of non-payment? That is, while these individuals appear to have the ability to pay \$250 today, will they have that same ability to pay an accumulated deferral balance of say, \$1,000 tomorrow?

**Table 4.2 Selected characteristics of deferrals versus those who fully pay within one year for individuals with beginning module balances of under \$250**

Feature	Deferral	Full Pay
Median Starting Balance	136	124
Median Income	19,181	22,592
Median Credit Balance <sup>1</sup>	10,445	10,472
Median Age	35	37
% Only One Module	57.8	81
% Filing Form 1040	26.9	44.5
% Filing Status of Single	62.8	56.2
% Remittance with Return	8.3	22.6
% Schedule C or F	5.9	12.1
% Homeowners	9	22.5
% With Investment Income	12.7	35.9

<sup>1</sup> Based on total outstanding credit as of 5/96.

**Section 4.1--Cases with Only One Module**

Of the 33,263 individuals in the panel with a starting balance exceeding \$250, 11,976 (45.4 percent) fully pay within one year. Of these, 77 percent have just one module. On the surface, it would seem that the number of modules is an important indicator of who is likely to pay. As it turns out, sophisticated classification techniques discussed later support this observation. Simple counting methods bear this out as well: individuals with just one module are 1.6 times more likely to fully pay in one year than not pay; those with multiple modules are 3 times more likely *not* to pay than pay. Intuitively, these simple odds would seem to suggest a separate investigation of cases with just one module.

To do so, several basic questions need attention: what characteristics, if any, distinguish those who pay from those who do not; what

proportion of these cases are repeaters; and are there other patterns not previously studied that can improve our understanding of this market?

Table 4.3 addresses the first issue by comparing those who fully pay within one year from those who do not (for those with a single module whose balance exceeds \$250). Patterns of risk are evident: those who fully pay are almost twice as likely to submit a remittance with their return; will have a module balance relative to income that is, on average, 50 percent lower than those who do not fully pay; and will be 1.5 times more likely to own a home or have investment income. Those who do not pay appear more likely to be single and slightly younger as well. One feature of particular significance that will be discussed in more detail below is the percentage of individuals who make at least one payment in 180 days: those who do so are roughly 3 times more likely to pay.

Before turning to that issue, there is the question of repeaters: of those cases with just one module, who had a delinquency on a balance-due return in the prior tax year? Of the 15,068 cases with one module whose balance was greater than \$250, 22.3 percent met this simple criterion. What is somewhat more interesting is that the repeater rate for those who fully pay within one year is higher (25.6 percent) than those who do not (17.3 percent). Also of interest: of those who fully pay and have a remittance with their return, 61 percent use a tax preparer. Are preparers correcting for noncompliance, for example, obliging a taxpayer to file a balance due return rather than deliberately overstate expenses or understate income to avoid owing taxes? Or are these simply corrections to an unanticipated liability based on a taxpayer's misunderstanding of tax law or return instructions?

Finally, Section 3 asked whether the timing of payments is important: do certain individuals have a higher probability of paying earlier than others? What inferences, if any, can be drawn from such a market? Table 4.4 addresses this question by looking at the distribution of payments within 180 days and the associated probability of full payment within 360 days. Selected financial and other characteristics are included in the hope of finding patterns that might also be useful for classification purposes.

Several insights can be drawn from Table 4.4. First, the obvious: those who delay payment (or possibly even contact), all other things equal, will be less likely to pay within one year (see Table 4.3). Second, there is evidence that those who delay payment are somewhat younger, file simpler returns, have smaller incomes, and are less likely to own a home or have investment income. But what other factors determine whether an individual delays their response to a notice? Could there be explanations that are systemic and treatable, such as poor toll-free telephone access or the ambiguity of a notice relative to industry billing statements?

Finally, an examination of account dispositions for those not fully paid at one year is warranted. Table 4.5 shows that almost 95 percent of these cases are in just three categories.

**Table 4.3 Selected characteristics of individuals with one module, by payment status<sup>1</sup>**

Feature	Pay	No Pay
Median Beginning Balance	\$800	\$1,015
Median Income	36,612	29,656
Median Balance Burden (%) <sup>2</sup>	2.2	3.4
Median Credit Balance	10,114	10,023
Median Credit Burden (%) <sup>3</sup>	2.8	3.4
Median Age	41	38
% Filing Form 1040	67.1	54
% Remittance with Return	40.7	23.7
% Who Use Preparer	51.7	43.6
% With Home or Investments	64.5	43.2
% Single Filers <sup>4</sup>	46.3	56.4
% With Schedule C or F	25.8	21.1
% Payment within 180 days	69.1	23.7

<sup>1</sup> Module balance greater than \$250.

<sup>2</sup> Starting module balance as a percent of income.

<sup>3</sup> Credit balance as a percent of income.

<sup>4</sup> Includes Head of Household and Widowers.

**Table 4.4 Distribution of individuals making at least one payment within 180 days, the probability of full payment at one year, and associated characteristics**

Number of Days from Assessment	% Making at least 1 Payment <sup>1</sup>	% Who Full Pay at 1 Year	Median Starting Balance	Median Income	Median Age	%1040 Filers	% With Home or Investments	% With Return Remittance
30 Days	6.4	77	\$967	\$38,928	44	72.8	68.7	62.1
60 Days	12.9	74.2	1,034	37,904	42	69.1	66.3	45.4
90 Days	6.4	60.8	943	34,826	40	61.3	57.4	31.7
120 Days	3.3	63.6	904	34,973	39	61.5	56.8	24.4
150 Days	3.1	57.8	730	30,643	38	55.8	49.5	24.5
180 Days	1.5	48.7	744	30,887	38	56.7	46.3	22.5

<sup>1</sup> See Table 3.3 for definition.

**Table 4.5 Selected characteristics of cases not fully paid one year after assessment**

Feature	Deferral Installment	ACS
Percent of Cases <sup>1</sup>	38.9	47.5
Median Balance	611	1,456
Median Income	23,755	35,504
Median Burden <sup>2</sup>	2.6	4.1
Median Age	36	40
% 1040	44.6	59.6
% Single	49.2	35.1
% Home/Invest	32.2	51.1
% Schedule C or F	16.1	23.5

<sup>1</sup> Percent of those not fully paid within one year.

<sup>2</sup> Median balance as a percent of median income.

There is evidence that cases in ACS present the greatest risk: they can be readily distinguished based on their high Median Burden—the balance owed as a percent of income. Who makes up this group? Not surprisingly, a significant percentage (54 percent) are those same individuals more likely to be young, single, and have no homeownership or investment income. As a result, a more detailed profile of this group is provided in Section 4.2.

It may also be of some interest that cases in deferral status have, on average, a module balance that is only slightly higher relative to income (2.6 percent) than individuals who fully pay within one year (2.2 percent; see Table 4.3). Because of this, and since they have a balance

owed on just one module, they may represent little or no risk. However, as discussed earlier, this may not be the case if the IRS is allowing consecutive deferrals. This question will be taken up again in Section 6.

#### Section 4.2--Classification and Risk

In Section 3, it was demonstrated that the number of modules is an important factor, all other things equal, in describing the likelihood of payment. It was that result which in fact motivated the work in this section; that is, of isolating and analyzing individuals with just one module.

Splitting the panel on this particular feature was no accident or guess, but rather the outcome of statistical, machine learning, and tree-based methods designed to look for such effects. Although a full description of these techniques is beyond the scope of this paper, several key results can be discussed.

First, individuals with just one module are far easier to describe and classify relative to the risk of non-payment than those with more than one module, as will be seen in the next section. In fact, every modeling technique mentioned above confirmed that only a few key features are needed to develop accurate risk analysis models for this group.

Second, and perhaps more importantly, such models can be examined heuristically through the use of decision trees to describe the rules involved, making them easy to interpret and

understand. Finally, in light of this, such models would be very easy to implement in practice.

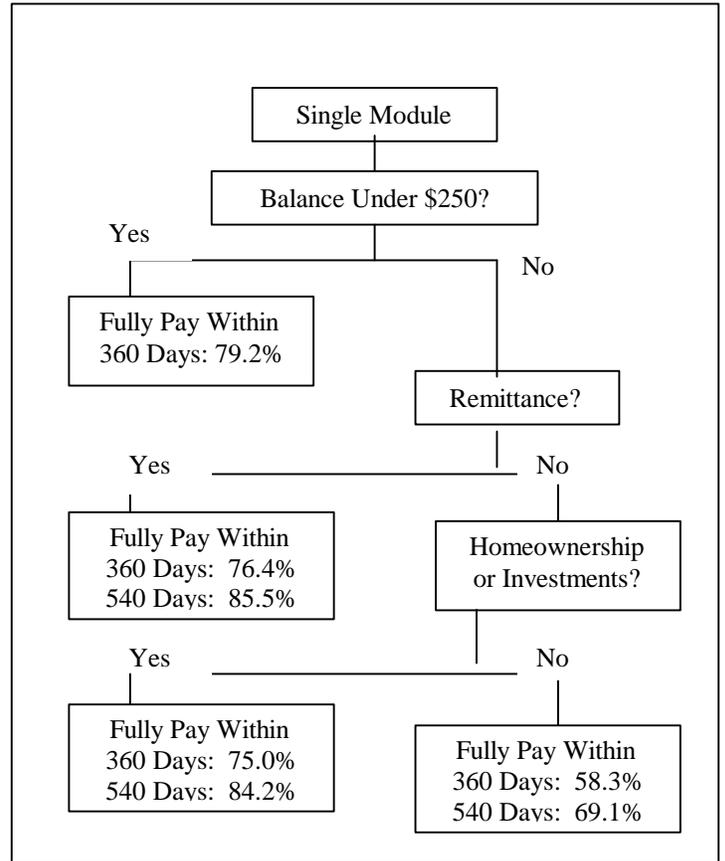
Combining the results of this section into such a heuristic decision tree gives Figure 4.1, which reveals four distinct subgroups: those with small balances, and those who: 1) have a remittance with their return; 2) have no return remittance and no homeownership or investment income; and 3) have no remittance but do have homeownership or investment income.

All other things equal, those who submit a remittance have a high probability of full payment within one year (76.4 percent), and might be considered low risk; the same is true for those who fail to remit but own a home or have investment income (75 percent). One group, however, has a distinctly different risk: those who have no return remittance and have no homeownership or investment income (less than 60 percent fully pay within one year).

Table 4.5 examines these last three subgroups in more detail, and it is not surprising that the same group of younger taxpayers more likely to delay payment (from Table 4.4) are those least likely to fully pay in one year—and more likely to be in ACS. Whether these results can be generalized for assessments from districts other than Maryland and D.C. is a testable hypothesis that warrants further analysis.

Several other hypotheses relating to this subgroup (those with no remittance and no homeownership or investment income) should be tested as well. First, this group has a significantly higher credit balance relative to income (43.4 percent), and there is the possibility that this burden, along with other debt not used in this analysis (for example, student loan payments), creates financial pressures too great for some. Second, is the repeater rate relatively higher for these individuals? (This question was not analyzed here.) Finally, given that they are younger and therefore more likely to change jobs, is this simply a withholding (W-4) problem?

**Figure 4.1 Example of a heuristic decision tree used to predict the likelihood of payment**



**Table 4.5 Selected characteristics of the subgroups represented in Figure 4.1**

Feature	No Remit (No Home/ Remit Investment)	No Remit (Home/ Investment)	
Median Balance	1,027	673	1,056
Median Income	40,900	24,223	43,279
Median Credit <sup>1</sup>	9,892	10,519	10,321
Credit Burden <sup>2</sup>	24.2	43.4	23.8
Median Age	42	35	44
% 1040	74.3	30.9	81
% Single	32.4	50	27.8
% Head House	8.3	19.1	12.6

<sup>1</sup> See Table 4.2 for definition.

<sup>2</sup> 'Median Credit' as a percent of median income.

**Section 5--High Risk Groups**

In the previous section, it was shown that the odds of not fully paying within one year were much lower for individuals with more than one module (3-to-1 versus those with just one module). Again, this makes sense if the individual is still paying on a prior year balance. However, this raises several simple but important questions for those with more than one module: 1) How much is being paid on prior module balances; 2) What factors can be used to predict the likely disposition of a new module; and 3) Is there a relationship between payment activity on prior modules and expected future payment on a new module?

Before exploring these questions, it is worth noting that individuals with multiple open modules can be studied in two ways: the characteristics of each module can be examined separately (module-level analysis), or in combination (entity-based analysis). While specific needs often drive which approach to take, any attempt to construct reliable models for classification purposes requires the use of both, and that is the direction taken here.

Of the 33,263 modules in this panel, 13,350 (40.1 percent) had more than one module (with the module balance on the new assessment exceeding \$250). Of these, 27 percent fully pay the most recent module within one year. Roughly 62 percent of those who fully pay their new module balance also fully pay their entity balance. For those who do not fully pay their new module balance, Table 5.1 examines the question of how much is being paid on both the module and entity balance over one year, by disposition of the account at the end of that year. Aside from those who fully pay, only those in an installment agreement are, on average, paying down their balances.

It is evident from the median balances reflected in Table 5.1 that individuals with multiple modules represent a significantly different risk than those with just a single module. This is true even for individuals in installment agreements, which one may be inclined to consider relatively less risky. This can be seen in Table 5.2, which shows that, for those in an installment agreement, the probability of making payments to any module is a decreasing function of the number of modules. That is, although the entity

balance is being paid down, the reduction is inversely related to the number of modules: those with two modules pay down, on average, almost 8 percent of their total entity balance over one year, while those with six pay down just over 2 percent.

**Table 5.1 Payments made on current module and entity balances over one year for cases with more than one module, by disposition**

Disposition	Starting Module Balance <sup>1</sup>	Ending Module Balance	Starting Entity Balance	Ending Entity Balance
Bankruptcy	\$1,418	1,384	4,764	4,899
Deferral	532	487	1,108	1,115
CNC	1,657	1,971	9,886	11,167
Installment	1,342	1,236	3,599	3,332
ACS	1,192	1,292	3,531	4,053
CFF	1,655	1,994	10,105	11,635

<sup>1</sup> All balance amounts are in medians.

**Table 5.2 Installment agreement payments over one year, by number of modules**

Number Of Modules	% <sup>1</sup>	Starting Module Balance <sup>2</sup>	Ending Module Balance	Starting Entity Balance	Ending Entity Balance
2	48.2	\$ 1,427	1,119	2,513	2,320
3	26.7	1,241	1,284	3,744	3,342
4	13.1	1,208	1,331	5,677	5,291
5	5.3	1,249	1,378	8,425	7,733
6	2.7	1,338	1,542	14,872	14,536
>6	3.6	1,682	1,984	34,969	35,286

<sup>1</sup> Percentage of installment agreement cases.

<sup>2</sup> All balance amounts are in medians.

Another troublesome feature of Table 5.2 is that for those with just two modules, the entity balance decreases by less than the current module balance. It seems that not only is the risk of non-payment in installment agreements

higher for those with more modules, but even for those with just two modules. Why is this? Accumulating interest alone? As it turns out, the reason can be attributed to new modules opening at the end of 360 days—to repeaters. In fact, of the 4,721 individuals in an installment agreement comprising Table 5.2, a full 2,232 (48.1 percent) had a new balance-due assessment the following year. The fact that nearly one out of every two individuals with multiple modules an installment agreement will be a repeater raises serious questions about the role the IRS should play in prevention management—especially for cases that may offer the greatest chance of recovery through treatment. (The question of whether a repeater rate for *this* panel is representative of the nation, however, is beyond the scope of this investigation.)

The next basic question that needs to be addressed is what features, if any, can be used to describe and predict the likely disposition of accounts. Table 5.1 clearly reveals that individuals in Bankruptcy, CNC, and CFF status have significantly higher entity balances relative to income. Is this feature alone sufficient for heuristically categorizing these individuals as inherently more risky? Probably not, although additional characteristics of these cases shown in Table 5.3 seem to support this conjecture: those in Bankruptcy, CNC, or CFF also have a higher number of modules, on average, than individuals in other dispositions; a higher relative accumulation of penalty and interest; and for bankruptcies in particular, relatively higher credit balances. In fact, if further research cross-validated the association between high credit balances and bankruptcies, the IRS might use such evidence to explore the impact of easy credit standards in the banking industry on potential loss of revenue to the U.S. Treasury. One additional feature of interest from Table 5.3 relates to individuals in CNC status: they tend to be nearly twice as likely, on average, to have a Schedule C or F than other groups. However, their average income, homeownership, and investment rates seem to reflect an unstable source of income. Clearly, this type of information—were it available at the time of assessment in the form of a risk analysis scoring system—might benefit collection activities in a wide variety of ways.

The last question to be explored here is whether a relationship exists between prior payment activity—or the disposition of a prior module—and expected future payments. For individuals with just one module, it was seen that timing played a key role: the sooner contact is made and payments started, the higher the probability of fully paying within a reasonable period of time. (Ironically, it may be timing alone that is responsible for a large percentage of multiple-module cases: if I begin payments late and don't have last year's balance paid off before being assessed with a new delinquency this year, my overall liability is compounded.) The question here is whether the disposition of the last module will be a determinant, all other things equal, of the likely disposition of a new module.

Table 5.4 investigates this issue by constructing the joint distribution of the probability that an individual will have a particular disposition for a new module given the current disposition of their most recent module. Of value here is the main diagonal of this table, which shows the probability of a particular disposition for a new module given that the last module is currently in the *same* disposition. For example, the probability that an individual's new module will be in bankruptcy given that their most recent module is currently in bankruptcy is 42.4 percent.

From this table, some useful results can be computed: the odds of full payment are 2 times greater for those whose prior module in either deferral or installment agreement than those with any other status; the odds of a new module being in either bankruptcy, CNC, or CFF status is over *11 times* higher for those whose prior module was in either of those three categories; and so on. In short, the joint probability distribution in Table 5.4—in combination with additional features from Table 5.3—appears to be very useful for developing reliable models that predict the likely disposition of multiple-module cases, to be discussed below.

**Table 5.3 Selected characteristics of multiple-module account dispositions at one year, by disposition category**

Disposition	Starting Module Balance <sup>1</sup>	Starting Entity Balance	Starting Penalty/Interest	Median Income	Median Credit Balance <sup>2</sup>	% With Schedule C or F	% 1040 Filers	% With Over 2 Modules	% With Home/Invest <sup>2</sup>	% With Return Remit <sup>2</sup>
Full Pay	\$ 912	2,198	172	35,933	9,867	24.0	61.4	28.8	53.9	22.5
Bankruptcy	1,418	4,764	639	38,091	12,127	28.6	72.0	65.6	60.8	11.1
Deferral	532	1,108	98	25,955	10,462	15.5	40.2	30.2	26.7	13.4
CNC	1,657	9,886	1,707	26,120	10,397	35.8	54.8	76.4	29.7	7.4
Installment	1,342	3,599	342	32,877	10,303	17.4	50.5	51.8	39.6	11.4
ACS	1,192	3,531	372	26,821	10,853	19.1	44.6	57.7	26.6	8.7
CFF	1,655	10,105	835	33,341	10,562	22.0	57.4	69.3	41.7	9.4

<sup>1</sup> All dollar amounts are in medians.

<sup>2</sup> See Table 4.3 for definition.

**Table 5.4 Joint probability distribution of the disposition of a new module at one year given the disposition of most recent module**

Disposition of New Module at One Year	Disposition of Last Module <sup>1</sup>							
	Percent of Cases <sup>2</sup>	Bankruptcy	Deferral	CNC	Installment	Full Pay	ACS	CFF
Bankruptcy	3.0	42.4	6.3	1.2	32.2	—	9.9	2.1
Deferral	9.5	0.4	56.5	8.7	24.0	—	4.3	0.4
CNC	5.1	1.1	5.0	28.0	20.9	—	24.3	4.1
Installment	42.7	1.1	12.8	0.8	61.0	—	13.3	0.9
Full Pay	24.9	1.9	22.2	1.7	53.6	—	9.9	1.2
ACS	9.8	3.4	24.1	2.9	30.5	—	25.5	0.8
CFF	2.3	1.7	10.9	0.4	13.4	—	19.7	26.4

<sup>1</sup> At the time of assessment of the new module.

<sup>2</sup> Represents the percentage of cases for new modules at the end of one year.

**Section 5.1--Classification and Risk**

Section 4 outlined a simple, heuristic approach for classifying individuals with one module based on just a few features. For example, almost 80 percent of those with one module who have a remittance with their return will fully pay in one year. Can a similar rule-based approach with such high accuracy be developed for individuals with more than one module—those representing significantly higher risk?

Unfortunately, very few heuristic rules could be found for this group: the information space is relatively more complicated, and thus the classification problem more difficult. However,

using statistical, machine learning, and tree-based techniques with many of the features covered in this section, risk assessment models can be developed with accuracy similar to those of the previous section (77 percent to 89 percent). There are several models to consider: 1) who is likely to fully pay their new module within one year—a strong indicator of who will fully pay their entity balance; 2) of those remaining, who is likely to be disposed in Bankruptcy, CNC, or CFF—cases with the highest relative risk; and 3) of those likely to enter an installment agreement, which individuals are most likely to make payments. Of course, there are other viable classification models as well, and those mentioned above are

by no means exhaustive. Those interested in a more detailed analysis of classification results from this study should contact the author.

### Section 6--Deferrals Revisited

Of the original 33,263 modules in this retrospective panel, 16,570 (49.8 percent) did not fully pay their balance within one year; of these, a full 30 percent were deferred (see Table 3.2), with the total amount deferred representing 7.7 percent of the total module balance remaining at one year. Of those in deferral at the end of one year, 35 percent had more than one module.

Earlier sections presented a cursory examination of deferrals and, based on those results, raised several critical questions about their relative risk: What is the rate of payment on deferrals over say, two or three years? What are the risks, if any, of deferring balances for individuals with more than one module? What is the probability that a deferral ends up in bankruptcy, ACS, or CFF? Do such probabilities depend on the number of modules?

To answer some of these questions, a separate panel of individuals from Maryland and D.C. was created using the methodology outlined in Section 2.1, with the only difference being the year of assessment—1995 instead of 1996. This earlier date will allow for retrospective tracking over three years instead of two. The new panel contained 27,305 cases, of which 14,508 (53.1 percent) did not fully pay within one year; of these, 3,922 (27 percent) were deferred. Of those in deferral at one year, 34.7 percent had more than one module.

At two years, only 1,491 (38 percent) of those originally deferred fully pay their balance; of these, 70 percent have just one module. Of those who do not pay, however, 61 percent have more than one module at two years. Table 6.1 shows that of those who did not fully pay within two years, only 79 percent remain in deferral status; 10 percent are in an installment agreement. However, almost 9 percent are in dispositions that might be considered categories of risk: bankruptcy, CNC, ACS, and CFF. More important is the percentage of originally deferred cases that have more than one module.

**Table 6.1 Distribution of deferrals after two years, by disposition and number of modules**

Disposition	% <sup>1</sup>	Number of Modules			
		1	2	3	4+
Bankruptcy	0.8	15	60	20	5
CNC	0.7	19	50	25	6.2
Deferral	79.3	46	31.5	16	6.5
Installment	10.1	11	35.6	38.1	15.8
ACS	6.4	12	34.6	32.7	21.2
CFF	0.7	13	56.3	12.5	18.7

<sup>1</sup> Percentage of all originally deferred cases that have not fully paid at two years.

At three years, 2,202 (56 percent) have fully paid their original balance, but as shown in Table 6.2, the proportion of cases in risk categories increases by 84 percent, from 9 percent to 16.6 percent. There is also a noticeable shift in the percentage of cases towards an even greater number of modules.

**Table 6.2 Distribution of deferrals after three years, by disposition and number of modules**

Disposition	% <sup>1</sup>	Number of Modules			
		1	2	3	4+
Bankruptcy	2	17.6	20.6	29.4	32.4
CNC	2.6	24.4	37.8	26.7	11.1
Deferral	72.4	48	29.1	14.0	8.9
Installment	9	10.6	19.4	35.6	34.4
ACS	10.5	13.3	27.1	30.9	28.7
CFF	1.5	4	20	44.0	32

<sup>1</sup> See Table 5.5.

These results are troubling for a number of reasons. First, Section 4 asked whether a high percentage of deferrals with balances of \$250 or less were, on average, being paid in two years—a more than reasonable amount of time

for such an amount under most financial conditions. If so, then these cases might truly represent zero risk. In fact, of the panel of deferrals above with beginning entity balances of \$250 or less, only 43 percent fully paid in two years.

Second, as with the issue of timing discussed in Section 4, one might infer from the above data that the probability of risk (i.e., of non-payment) for deferrals is an increasing function of time: the longer the deferral is ignored, the greater the chances of it moving into a bankruptcy, ACS, or CFF status—all costly to the IRS.

Finally, there is strong evidence from the above tables of consecutive deferrals, which by definition increases the number of modules and hence entity balance as well. Allowing a new deferral on top of an old one may be compounding the risk of non-payment: as the number of modules increases, the accumulation in entity balance may eventually become unmanageable relative to a fixed income source.

### Section 8--Discussion and Conclusion

One need only examine Table 7.1 to get a renewed sense of the challenges inherent in delinquent individual collections. For the panel of 27,305 cases studied in Section 6 with new assessments in May of 1995, the total beginning entity balance was just over

**Table 7.1 Payments on aggregate entity balances over three years**

Date	# of Entities <sup>1</sup>	Percent Change <sup>2</sup>	Entity Balance <sup>3</sup>	Percent Change
May-95	27,302	-	\$134,319	-
May-96	19,061	30.2	125,902	6.3
May-97	13,978	26.7	121,531	3.5
May-98	10,918	21.9	120,507	1.0

<sup>1</sup> Original number of entities is shown at 5/95.

<sup>2</sup> Percent reduction from previous year.

<sup>3</sup> In millions.

\$134 million. After one year, 6.3 percent was paid; after two years, an additional 3.5 percent was paid; after 3 years, just 1 percent. If one were to track this aggregate payment function beyond three years, it may show continued

payments, but at a decreasing rate. It would also show dispositions for these cases in roughly the same proportion as presented in Section 5.

Who are these individuals remaining after three years with large entity balances that are at risk of non-payment? What factors can be used to describe such risk? Clearly, this research provides concrete answers to these basic questions. Can we identify these cases early enough so that, where appropriate, different strategies or treatments can be pursued? The analysis presented in Sections 4 and 5 would indicate that the answer is yes.

In short, the results of this study suggest that the ability to identify and distinguish individuals who are likely to pay from those that are not—in a framework of relative risk at the time of assessment—may be an important ingredient for helping reduce the inventory reflected in Table 7.1. It also demonstrates that building accurate models for such purposes is not beyond the reach of the IRS. At a minimum, a system designed to score delinquent accounts at the time of assessment could offer:

- Greater ability to identify and devote resources to high-risk accounts before they become a financial burden.
- Potential to reduce intrusiveness on taxpayers who are otherwise likely to pay.
- Ability to prioritize workload for more efficient utilization of IRS resources.

Before the merits of developing and implementing such a system are debated, the work of this paper must be cross validated against other district office data and time periods. However, if the patterns are favorably close, it would imply that a new risk analysis system such as that proposed should be given serious consideration—perhaps for businesses as well as individuals. It would also present an opportunity for additional research to study unique customer markets found in Sections 4 and 5. For example, why is the repeater rate for individuals in an installment agreement so high? And why are such a high percentage of low-risk delinquencies (Section 4) associated with returns prepared by tax practitioners?

Such new research might require a more qualitative framework—perhaps through the use of survey instruments or focus groups—to investigate such questions. Based on inferences

from this paper, it might also address, with monetary quantification where possible, several other topics:

- Certain taxpayers demonstrate a higher propensity for delaying payment on a notice. Would such taxpayers respond sooner to a billing statement such as those used by utilities, banks, and other lenders in industry? Such a bill might show a “Minimum Amount Due”, in effect permitting either full payment or automatic installment.
- Is the probability of delaying payment related to poor toll-free telephone access rates? If a taxpayer calls about a notice but is unable to speak to a customer service representative, what percentage abandon their efforts as a result?
- The repeater rate for multiple-module cases in an installment agreement is nearly 50 percent. If states have the authority to mandate “Driver Education” classes for those with poor driving records, could the IRS secure the authority to mandate a similar program? Would such tax counseling reduce repeater rates?
- Individuals who tend to be younger and single also tend to have higher credit balances relative to their incomes. Would partnership programs with local financial firms help counsel these individuals through education or credit consolidation?
- If individuals are 30, 60, or 90 days late on making payments through a newer billing mechanism, could it be reported to a credit bureau? What benefits might accrue from such action?
- About two-thirds of individuals who will likely pay their delinquent balance also use a tax preparer. What percentage of returns are being corrected for an anticipated tax liability based on a misunderstanding of tax law or return instructions? That is, to what extent does tax law complexity contribute systemically to delinquencies?

- Payment rates on deferrals are much lower than expected two or three years after assessment. One reason for this may be that the IRS allows new deferrals on top of old ones, creating the potential for balances that grow too large relative to income. Could the IRS prevent this by simply not allowing consecutive deferrals?
- There is evidence that payment rates on deferrals are inversely related to the time the case remains in deferral. Could the IRS reduce its exposure to this risk by routing deferrals to ACS after a specific period of time, say one year?

It is hoped that this analysis, as well as results from any future research proposed above, provides more than just a new look at an old problem. Given the magnitude and direction of aggregate delinquent balances, it would seem that much more is needed. A modern system for accurately identifying risk at the time of assessment might just be a good place to start.

#### **Acknowledgments**

The author wishes to acknowledge the Compliance Data Warehouse, without which this research would not be possible; the support of Maury Harwood, Rudy Estrada, and Bruce Colton; and the helpful comments of Mike Graeber, George Coakley, Doug Beazley, and Joel Friedman.

---

#### **About the Author(s):**

**Jeff Butler** is an Operations Research Analyst in the AC (Research and Statistics of Income) area. He received his M.A in Economics in 1988 from Penn State University. He has been with the IRS since 1988.

# Application of Scientific and Statistical Methods to an Operational Program: A Case Study of the Bank, Post Office and Library Program

By Denise York Young and  
Erika D. Alexander

*Prior to 1991, the IRS recorded very limited data on individual tax products delivered through the Bank, Post Office, and Library (BPOL) Program. Continual efforts since then to improve both the quantity and quality of data have enabled more systematic research on ways to improve the program. As described in more detail in this article, the use of quantitative methods has led to major modifications to the BPOL Program. The modifications have served to reduce the number of forms remaining at BPOL outlets at the end of the filing season, while increasing the number of taxpayers filing forms obtained from BPOL outlets. The use of scientific and statistical methods in the BPOL Program demonstrates a model that can be applied to other operational programs. The primary components of this model are (1) initiation of data collection processes, (2) implementation of data-driven operational decisions, and (3) control of external factors.*

## Introduction

The Internal Revenue Service's Bank, Post Office, and Library (BPOL) Program distributes Federal tax forms to the public through approximately 45,000 outlets nationwide. A BPOL account is a bank, post office, library, or other entity with a formal arrangement with the IRS to distribute tax forms through the BPOL Program. A BPOL account consists of one or more outlets. The BPOL Program serves over 18 million taxpayers, and has a printing budget of \$7 million. In 1991, in an effort to improve its efficiency, the IRS began studying this program using a variety of scientific techniques and statistical methods. This work has resulted in the following improvements from 1991 to 1997:

- an increase in the number of taxpayers filing forms obtained at BPOL outlets (from 13 million to 18 million),
- a decrease in the number of surplus forms remaining at BPOL outlets at the end of a filing season (from 77 million to 43 million),
- a decrease in the number of forms sent to BPOL outlets (from 394 million to 361 million), and
- a decrease in the number of BPOL locations distributing forms (from 100,000 to 45,000).

There are 10 standard items that all BPOL accounts receive: Form 1040, Form 1040A, Form 1040EZ, Schedule A&B, Schedule EIC, Schedule 1, Schedule 2, Instructions 1040, Instructions 1040A, and Instructions 1040EZ. The improvement efforts described in this paper were

aimed at all 10 products, but particularly the Form 1040. Tables 1, 2, and 3 contain historical data relative to Forms 1040, 1040A, and 1040EZ, respectively. They show the volumes distributed to BPOL outlets, the surplus amounts remaining at the end of each filing season, and the numbers of taxpayers filing the forms obtained from BPOL outlets.

As a result of the research, the number of Forms 1040 distributed to BPOL accounts decreased about 25 percent since the 1992 filing season, while the number of taxpayers filing a Form 1040 obtained from a BPOL outlet increased. Moreover, the number of surplus Forms 1040 remaining in BPOL outlets at the end of the filing season was reduced by almost half. There were no reductions in the number of Forms 1040A and 1040EZ sent to BPOL outlets; however, their surpluses at the end of the filing season also decreased.

## Data Collection Processes

Prior to 1991, the IRS recorded only the most basic data needed for delivering forms to the outlets. The program recorded the type of account (i.e., bank, post office, library, other) and the number of locations a particular account serviced. But this was often inaccurate and incomplete. Ordering information for specific accounts was not retained from year to year, making analysis of historical trends at a micro level impossible. Even though the Service knew the quantity of forms sent to BPOL outlets, it did

not know the portion actually filed by taxpayers, or the unused portion at the end of the filing season.

In order to increase efficiency, the BPOL Program committed itself to increasing the quality and quantity of its operational data. For instance, to gain insight on taxpayer usage and end-of-season surpluses, the IRS began printing source codes on the forms distributed by BPOL outlets in 1991 (i.e., "B" for bank, "L" for library, and "P" for post office). From an operational standpoint, source codes added complexity to the BPOL Program. The source codes had to be added manually to the proofs of each tax form, additional contracts had to be administered, and the distribution center had to handle additional products. However, the data collection aspects of this initiative were made easy by an agreement reached between the BPOL Program and the IRS's Statistics of Income (SOI) Division. The SOI Division conducts an Early Tax Estimates (ETE) Study that draws a random sample of approximately 20,000 returns each year. As a result of the agreement, SOI modified its ETE Study to capture source code information used to estimate the lower bound on the number of people served by BPOL outlets. It should be noted, though, the ETE Study does not capture data on an undetermined (and probably large) number of taxpayers that picks up forms and instructions at BPOL outlets but does not file them.

To supplement the information provided by the ETE Study, the BPOL Program established an annual inventory report in April 1993 to determine the end-of-filing-season surplus of the 10 standard products at each BPOL outlet. Because of the expense of mailing the inventory report to all accounts and the time burden placed on respondents, a stratified sample was used to obtain information on accounts that received large quantities of forms. This sampling method also randomly selected accounts in other size categories, as well as estimated the total amount of surplus for each of the 10 products. Over time, the sampling process was refined to reduce the sampling error, by increasing both the sample size and the response rate. In 1993, 6,414 inventory reports were mailed and 2,744 were returned with usable data. By 1997, the inventory report was mailed to 10,088 accounts and usable data were obtained from 7,292 of them.

### **Implementation of Data-Driven Operational Decisions**

The process of integrating data into BPOL operational decisions has evolved since 1991. At the onset, much of the data was either unavailable or unreliable; however, the quantity and quality of data for making operational and managerial decisions have increased greatly over time.

#### Development of Recommended Amounts

Statistical methods first were introduced to the BPOL Program to develop recommended delivery amounts for each account for the 1992 filing season. Analysis of historical operational data suggested the number of forms distributed to BPOL accounts could be reduced. The ratio of number of Forms 1040 available in BPOL outlets to total number of individual tax returns filed was developed as a measure of abundance of forms in an area. This ratio was computed for each county and metropolitan statistical area in the country. Accounts in those counties and metropolitan statistical areas whose ratio exceeded the median by a given amount had an adjustment factor applied to the total amount of forms received the previous year. For example, the adjustment factor initially was applied to all accounts in counties and metropolitan statistical areas whose ratio was more than 1.5 times the median ratio. This methodology was applied incrementally over time, reaching full implementation in the 1996 filing season when the adjustment factor was applied to all accounts in counties and metropolitan statistical areas whose ratio was above the median. The use of this type of adjustment factor resulted in recommended amounts that were less than or equal to the total amount of forms received by an account the previous year. Depending on the year, the percentage of accounts whose recommendation was reduced by this process ranged from less than 1 percent to over 28 percent.

The BPOL Program further modified the process of developing recommended amounts in the 1994 filing season, by incorporating data from the inventory report. Prior to that, BPOL accounts whose amounts were not adjusted by the statistical ratio method (described in the preceding paragraph) received the same recommended amount as the previous year (i.e., initial amounts plus any resupply amounts). This was because no information existed to indicate

whether an account ran high or low on forms. For those accounts that returned an inventory report, the number of forms and instructions sent could be reduced or increased in response to reported surpluses and shortages. The use of this method, in conjunction with the Form 1040 ratio measures, generated annual savings of approximately \$750,000 over the prior method of developing recommended amounts.

#### Computation of Plan Ratios

The BPOL Program used data from the inventory report to compute plan ratios as another way to improve efficiency. Plans are groups of the 10 standard items packaged together in specific quantities. Prior to the 1994 filing season, the number and mix of items in a given plan were not based on quantitative analysis, but on past experience. Beginning in the 1994 filing season, statistical models were developed based on data from inventory reports.

Given the inventory report records the quantity of standard items that remain after the filing season, product usage was obtained by subtracting amounts remaining at the end of the filing season from total amounts sent to each account. Usage trends were developed and adjustments were made to each of the plan sizes to better reflect the observed usage patterns. By creating plans that more accurately reflected accounts' needs, the BPOL Program reduced both surpluses and shortages. The data analysis provided a sound method to determine product quantity in a plan.

#### Recommendation of Resupply Amounts

Once instituted, the dataset from the inventory report was used in a variety of ways -- some of which had not been envisioned at the onset. For example, matching inventory report information with ordering data revealed BPOL accounts that reordered forms throughout the filing season were, in fact, more likely to have surplus forms left at the end of the filing season. This discovery led to the development of suggested resupply amounts, which decreased in accordance with ordering trends as the filing season neared end. Institution of these tables of resupply amounts helped reduce the number of forms remaining at BPOL outlets at the end of the filing season.

#### Modification of Number of Outlets

The efforts described thus far concentrate on the use of quantitative methods to modify the number of *forms* available at BPOL outlets. Statistical methods also were used to identify counties either in need of additional BPOL *outlets* or counties saturated with them. The BPOL Program used a variety of measures and data sources to identify such counties. For example, data from the three IRS Distribution Centers were analyzed to determine which counties had a high number of orders for tax forms on a per capita basis. In addition, ratios of the number of BPOL outlets to number of tax returns and the number of forms available in BPOL outlets to the number of tax returns were computed for each county. The BPOL Program then focused recruitment efforts for outlets in counties with values in the bottom 10 percent for such measures. All but 18 of the nation's 3,140 counties had at least one BPOL outlet for the 1997 filing season.

Counties with an abundance of BPOL outlets experienced limited additional growth. For example, in the 1997 filing season a county was identified as saturated if it had more than 25 outlets *and* had more than 4.8 outlets per 100 square miles. These values were derived from statistical analysis of the data, and represented values that were more than 1.5 interquartile ranges above the 75<sup>th</sup> percentile (which is a commonly accepted statistical measure of extreme values). Approximately 200 of the nation's 3,140 counties fit this definition of saturation for the 1997 filing season.

#### Removal of Banks as Distribution Outlets

Removal of most banks as distribution outlets was perhaps the most visible result of the BPOL Program analyses. Analysis of the source code data from the ETE Study and the analysis of BPOL data over several years led to the conclusion banks were the most ineffective of the three major outlet types. With the IRS Commissioner's approval, in 1996 the BPOL Program eliminated most banks as distributors of tax forms and information. Statistical analysis demonstrated no adverse effect on taxpayers or the IRS as a result of this action. There were fewer BPOL outlets and fewer forms sent to BPOL outlets, yet more returns filed with BPOL source codes (refer to Tables 1, 2, and 3). Furthermore, there was no appreciable increase

in the number of forms obtained at IRS offices nor in the number of orders placed into IRS distribution centers. The BPOL Program successfully moved taxpayers from banks to post offices and libraries with little negative impact.

**Organizational Changes/External Factors Affecting the BPOL Program**

In spite of statisticians' best efforts to control external variables, an operational program can not exist in a completely controlled environment. Over the years there has been a variety of uncontrolled factors, both external and internal to the BPOL Program, that has affected the ability to make and measure changes to the program. Budgetary restrictions limited the initial shipment of forms to the BPOL accounts for the 1997 filing season. Moreover, accounts were unable to change the recommended amounts, and order more forms. When the budgetary restrictions were lifted later in the 1997 filing season, reorder amounts were extremely high, partly as a result of over-reaction to the initial restrictions.

At the same time the BPOL Program was striving to increase efficiency, additional demand was placed on it by the decreased availability of forms in other places. The number of IRS offices distributing tax forms decreased from 625 for the 1991 filing season to 507 for the 1997 filing season. As a result of the budgetary restrictions for the 1997 filing season, 46 million taxpayers that filed a practitioner-prepared return the previous year did not receive a postcard or tax package from the IRS. For that same filing season, approximately 22 million taxpayers received 1040EZ Telefile packages that did not contain the traditional paper Form 1040EZ and Instructions 1040EZ. (Since only 5 million filers of Form 1040EZ used Telefile, we estimate 17 million obtained paper forms elsewhere.) The exclusion of these forms from tax packages added over 3 million taxpayers to the BPOL Program.

**Conclusions**

IRS's experiences with the BPOL Program since 1991 demonstrate that the application of scientific and statistical methods to an

operational program provides opportunities for more effective management, which can lead to a more efficient program. However, such integration requires commitment from management and workers. Resources must be dedicated to increasing both the quantity and quality of data available for statistically based decision-making. Conflicts sometimes occur between the allocation of resources for "operations" versus "research." In reality this dichotomy is artificial because the purpose of "research" is to improve the efficiency of the operational program. Above all, it is important to measure the results of new methods and compare them to previous ways of doing business. In addition, efforts must be made to limit the impact of external forces in order to measure results in a meaningful and objective manner.

---

**About the Author(s):**

**Denise York Young** worked as a Mathematical Statistician in the Tax Forms Marketing Analysis Program of the IRS's Multimedia Production Division from 1990 to 1998. Currently, she is Assistant Director of Strategic Planning and Analysis at the University of Texas at Dallas. She received her M.S. degree in Applied Statistics and Animal Genetics from North Carolina State University and is enrolled in the Ph.D. program in Higher Education and Educational Research at the University of North Texas.

**Erika Alexander** is a Mathematical Statistician in the Tax Forms Marketing Analysis Program of the IRS's Multimedia Production Division. She received her M.S. degree in Statistics in 1991 from Texas A&M University and is currently a doctoral candidate at the University of North Texas in Educational Research and Computer Education/Cognitive Systems. She has been with the IRS since 1992.

**Table 1 Historical Data for Form 1040 (in Millions)**

Filing Season	Amount Sent	Amount Remaining	Number Filed with BPOL Source Code
1991	96.5	n.a.	4.7
1992	101.1	n.a.	5.1
1993	91.8	14.8	5.0
1994	86.4	14.2	5.2
1995	81.0	12.0	5.3
1996	67.6	8.2	6.0
1997	73.3	7.6	6.4

**Table 2 Historical Data for Form 1040A (in Millions)**

Filing Season	Amount Sent	Amount Remaining	Number Filed with BPOL Source Code
1991	54.8	n.a.	3.0
1992	58.3	n.a.	4.0
1993	71.3	11.0	3.5
1994	81.2	16.8	3.6
1995	75.7	15.0	4.0
1996	62.5	10.0	3.8
1997	66.5	9.3	3.9

**Table 3 Historical Data for Form 1040EZ (in Millions)**

Filing Season	Amount Sent	Amount Remaining	Number Filed with BPOL Source Code
1991	53.5	n.a.	4.0
1992	57.1	n.a.	4.8
1993	70.4	11.5	4.5
1994	80.5	17.0	5.2
1995	74.7	15.4	5.0
1996	60.8	9.7	5.1
1997	65.6	6.9	7.8

## Using Data Reduction Techniques to Analyze Baseline Profiles

By Larry May and  
Anne Steuer

*The Baselining National Compliance Measures project generated individual tables consisting of eleven compliance measures across three years for each of 71 broad taxpayer market segments for 33 IRS district offices. Analyzing this copious output presented a considerable challenge. Data reduction techniques were explored as an alternative to commonly used ad hoc methods of interpretation. Factor analysis together with cluster analysis was used to identify associations between the eleven measures and among the broad market segments. Preliminary results demonstrate that these various measures can be reduced into the three broad underlying compliance dimensions --market size, balance due, and accuracy. Further, market segments can be clustered into relatively homogenous groups with respect to these broad compliance dimensions and "compliance profiles" generated for each cluster. These results can be used to identify and target market segments with similar compliance characteristics.*

### Introduction

An expectation of the fiscal year (FY) 1997 Research Plan was to identify major non-compliant market segments<sup>24</sup> through baselining accuracy, timeliness, and payment of National, Regional, and District populations. The project team considered several sources in determining which market segments and measures to use. These included, but were not limited to, the segments and measures used in the FY 1996 Research Plan baseline objective, feedback from the members of the Market Segmentation and Profiling Cooperative Strategy Working Group, District Office Research and Analysis (DORA) Chiefs and their staffs, the FY 1997 Annual Compliance Plan and the FY 1997 Research Plan. The source of data for this project was a subset of the full 1040 Filers Model of the Compliance Research Information System (CRIS) known as CRIS-Lite. Twenty-six market categories were identified in the original plan.

These 26 categories further were subdivided into 71 market segments for evaluation, including the total population for comparison purposes. District baseline tables were

generated, presenting eleven compliance measures for each of the 71 market segments.

<sup>24</sup> Market segments are groups of taxpayers with some common characteristic(s) such as all taxpayers claiming a refund, all businesses in the same industry, or all taxpayers with a Schedule A, for example.

With 33 districts, there were a total of 2,343 tables that needed to be analyzed for a national perspective. Analyzing this copious output presented a considerable challenge. Data reduction techniques were explored as an alternative to commonly used ad hoc methods of interpretation. Factor analysis together with cluster analysis was used to identify associations between the eleven measures and among the broad market segments.

### Data Reduction Methodology

The eleven measures and their original groupings are listed below:

#### Market Size

- Estimated Population

#### Timely Payment

- Total Unpaid Tax at Time of Filing
- Average Unpaid Tax at Time of Filing
- Percent Dollars Unpaid at Time of Filing
- Percent Returns Unpaid at Time of Filing

#### Timely Filing

- Percent Returns Late

#### Tax Accuracy

- Total Taxes Reported
- Total Predicted Tax Increase (PTI)
- Average Predicted Tax Increase
- Voluntary Compliance Level (VCL)
- Estimated Percentage of Returns Accurately Filed

Some of the measures assess the same compliance aspect; for example, *Percent of Dollars Unpaid at Time of Filing* and *Percent of Returns Unpaid at Time of Filing*. Other measures even have a direct mathematical relationship: *Total Unpaid Tax at Time of Filing* and *Average Unpaid Tax at Time of Filing* (the average is simply the total divided by the population). All four of these measures are slightly different evaluations of the same attribute: timely payment.

Relationships between some measures and attributes, or characteristics, are not as straight forward. For example, individuals who owe taxes may have a greater tendency to submit a less accurate return in an attempt to reduce their balance due. Accordingly, there could be a relationship between the four “payment” measures and the characteristic of “accuracy”.

In a traditional interpretation of these eleven measures two barriers are presented. First, how should we compare two market segments where the various measures are presenting a mixed message — market segment A has a higher *Percent of Returns Unpaid at Time of Filing* while market segment B has a higher *Percent of Dollars Unpaid at Time of Filing*. Second, it is difficult and subjective to infer the indirect relationships. Exactly how much of an influence does *Average Unpaid Tax at Time of Filing* have on accuracy? Factor analysis overcomes both of these barriers. Factor analysis creates characteristic scores that incorporate the influence of all the measures including the weak and indirect relationships.

## Factor Analysis

### Description and Methodology

Factor analysis explores the relationships between all the measures and attempts to identify the underlying characteristics. Having identified them, it quantifies the association between each measure and each characteristic. Measures with strong associations to a characteristic contribute more to determining the final characteristic score; this is referred to as loading high. Other measures with weak associations load low. Each measure can influence (positively or negatively) each individual characteristic. Usually a measure will load high on one particular characteristic and

load low to moderate on the other characteristics.

In our factor analysis each market segment was regarded as an observation with eleven associated measures. A principle components factor analysis was run to identify underlying characteristics that explain the correlations among the set of measures. Its purpose is to summarize a large number of variables or measures with a smaller number of characteristics while preserving as much of the total sampling variation, and thus original information, as possible. For this analysis, each individual baseline table is regarded as a multivariate observation. There are 71 tables per DORA; thus with 33 DORAs there are a total of 2,343 tables nationally. Each table has eleven measures that jointly describe compliance behavior for that particular market segment.

Refer to Appendix A for a more complete description of factor analysis, as it is used in this work.

### Results

Initially the correlation matrix of the eleven measures was examined to assess the suitability of the data for factor analysis. Of 55 correlations, 23 are greater than 0.30 suggesting factor analysis is appropriate for these data.<sup>25</sup> Further, the Barlett Test of Sphericity is significant at a level less than 0.0001<sup>26</sup>. The Kaiser-Meyer-Olkin Overall Measure of Sampling Adequacy is 0.71034, which is within the acceptable range.<sup>27</sup> The measures of sampling adequacy for all variables except VCL

<sup>25</sup> Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. 1995. *Multivariate Data Analysis*, 4th edition. Prentice-Hall, Inc. The authors suggest the initial general rule of looking for a substantial number of correlations greater than 0.30 for applying factor analysis.

<sup>26</sup> This test is used to assess the overall significance of the correlation matrix -- that is the off diagonals are nonzero. Statistically significant results indicate that the null hypothesis of zero correlations can be rejected. The results of this test indicate that the non-zero correlations in the correlation matrix are most likely not due to random chance. This test, like most statistical tests is sensitive to sample size. The larger the sample size, the greater the ability of the test to detect smaller departures from zero. Thus it is also important to examine the magnitudes of the correlations.

<sup>27</sup> The Kaiser-Meyer-Olkin Measure of Sampling Adequacy is an index ranging from zero to one. Values under 0.50 are considered unacceptable. A value of one indicates that each variable can be predicted perfectly using the other variables.

are greater than 0.50. Thus, only VCL falls in the unacceptable range, suggesting this measure has little association or relationship to the remaining ten measures. VCL was omitted from the data set and the correlation matrices and associated tests recomputed.

Under the revised data set, 18 of 45 correlations are greater than 0.30 and the Kaiser-Meyer-Olkin overall measure of sampling adequacy is slightly larger. As before, the Barlett Test of Sphericity is significant at a level less than 0.0001. All the variables in the reduced data set meet the individual measure of sampling adequacy threshold. Based on these results, the reduced data set was used for the factor analysis.<sup>28</sup>

Table 1 gives the results for the rotated solution.<sup>29</sup> The columns listed -- Factor 1, Factor 2, and Factor 3 -- give the factor loadings for each variable and the communality column provides a measure of the degree to which each measure is "explained" using the three factors.<sup>30</sup> All the measures have a high communality with the exception of *Percent Returns Late*. For this measure, the three factors account for only approximately a third of its variability. The sum of squares and percentage of trace indicate the relative importance of each factor in accounting for the variability of the eleven measures.

Four measures, *Estimated Population*, *Total PTI*, *Total Tax Due* and *Total Tax Dollars Unpaid* at filing load highly for Factor 1 and all other measures have relatively low loads. This suggests Factor 1 characterizes market segment size. The four measures, *Average Tax Dollars Unpaid* at filing, *Percent Tax Dollars Unpaid* at filing, *Percent Returns Late*, and *Percent Returns Unpaid* load highly for Factor 2. This suggests Factor 2 characterizes payment. The two measures, *Average PTI* and *Percent Returns Accurately Filed* load highly and in reverse directions for Factor 3. *Average PTI* loads positively and *Percent Returns Accurately Filed* loads negatively. This suggests Factor 3 characterizes accuracy. Note, *Average Tax*

<sup>28</sup> For computational details see Internal Revenue Service North Central Office of Research and Analysis, 1997. Baseline CRIS-Lite RP97-1.02 Volume 2 Appendix B of the National Report.

<sup>29</sup> Ibid.

<sup>30</sup> The factor loadings give the correlation between the variables and the factors. Communality ranges from 0 to 1 with 0 indicating no variation is explained by the factors to 1 where all the variation is explained by the factors.

*Dollars Unpaid* at filing loads moderately on both the payment and accuracy factors. Factor 1 (market segment size) accounts for 34.4 percent of total variability, Factor 2 (payment) accounts for 24.8 percent, and Factor 3 (accuracy) accounts for 21.2 percent.

## Cluster Analysis

### Description and Methodology

The other data reduction technique considered was cluster analysis. With limited resources, the IRS must attempt to address issues of non-compliance with the broadest possible wholesale approach. If Retail Food and Beverage has the same compliance levels as Hotels and Lodging it may be more appropriate to view these segments as sub-components of a larger market segment. Accordingly, the 71 market segments were subjected to cluster analysis in an attempt to identify market segments with common compliance characteristics.

Cluster analysis is a mathematical technique in which the difference between objects, relative to some attribute or set of attributes, is quantified. Similar objects are grouped together to form a collective object, or cluster. A successful cluster analysis will take a large number of observations and classify them into meaningful groups with minimal loss of information.

As with factor analysis, one of the primary objectives of cluster analysis is the reduction of data to aid in the interpretation of results. The term cluster analysis actually refers to a number of different techniques that all attempt to classify observations according to their common relationships. These techniques primarily differ in the way they measure similarity or difference and how they group the objects together.

After identifying the underlying characteristics from the eleven measures, focus was directed towards the 71 market segments that made up the baseline study. The motivation for segregating the population into these market segments was to answer specific questions posed during the planning phase. The basis for many of these questions was exploratory in nature. Also, the composition of the market segments was not mutually exclusive; a taxpayer may belong to two or more segments.

For example, a taxpayer could be a member of the Paper Filing market segment as well as a member of the Self Prepared market segment. With regard to compliance characteristics, were

these market segments truly different or were they different “views” of substantially the same market?

**Table 1. Rotated Component Analysis of Three Factors**

Variables		Factor 1 Size	Factor 2 Payment	Factor 3 Accuracy	Communality
X <sub>1</sub>	Estimated Population.	0.84	-0.19	-0.38	0.89
X <sub>8</sub>	Total PTI	0.95	-0.16	-0.14	0.94
X <sub>9</sub>	Total Tax Due	0.95	-0.18	-0.10	0.93
X <sub>10</sub>	Total Tax \$ Unpaid	0.86	0.31	-0.01	0.83
X <sub>2</sub>	Average Tax \$ Unpaid	0.12	0.69	0.54	0.78
X <sub>3</sub>	Percent Tax \$ Unpaid	0.08	0.91	0.13	0.85
X <sub>4</sub>	Percent Returns Late	-0.27	0.51	-0.04	0.33
X <sub>7</sub>	Percent Returns Unpaid	-0.09	0.85	0.11	0.73
X <sub>5</sub>	Pct, Rtns. Accurately Filed	0.27	-0.12	-0.87	0.83
X <sub>6</sub>	Average PTI	-0.14	0.08	0.93	0.90
Sum of Squares (eigenvalues)		3.44	2.48	2.12	<u>Total</u> 8.04
Percentage of Trace*		34.4	24.8	21.2	80.4

Trace = 8.03747 (sum of the eigenvalues)

Refer to Appendix B for further description of cluster analysis.

Results

Cluster membership of each market segment was obtained for cluster solutions from 2 clusters through 20 clusters. At each stage, an existing cluster is segmented into two groups.<sup>31</sup> In some cases the new clusters have the same compliance attributes as the “parent” cluster (a compliant cluster yielding two compliant clusters or a non-compliant cluster resulting in two non-compliant clusters). In these situations, the cluster was probably split due to the market size characteristic. In other cases non-compliance was further distinguished because the two emerging groups are different: one compliant and one non-compliant. For selected cluster solutions, additional inquiry was conducted,

<sup>31</sup> Since our objective was to identify the cluster solution with the minimum number of groups, our analysis evaluated the changes in cluster groupings from the 2-cluster solution upwards to the 20-cluster solution — a divisive analysis. This is in contrast to the way the clusters were actually built; from 20 down to 2 — an agglomerative procedure.

including a cross-tabulation of cluster membership and market segment. The cross-tabulation identified which market segments are associated with which clusters and permitted the assessment of optimization criteria two and three.

The 17-cluster solution presented seven non-compliant groups: clusters 10, 11, 13, 14, 15, 16 & 17. The population of the market segments included in cluster 14 was very small and this cluster was ignored. This was deemed to be the optimum cluster solution based on the criteria previously defined.

Figure 1 is a condensed dendrogram of the 17-cluster solution. This shows the relative similarity of each of the seventeen clusters. Clusters 14 and 16 are the most similar, since they have the shortest horizontal lines prior to joining together. These two clusters were the ones separated when moving from the 16-cluster solution to the 17-cluster solution. This resulted in the small population market segments of cluster 14 being split from the balance of cluster 16.

**Figure 1: Condensed Dendrogram using Ward's Method — 17 Clusters Shown**

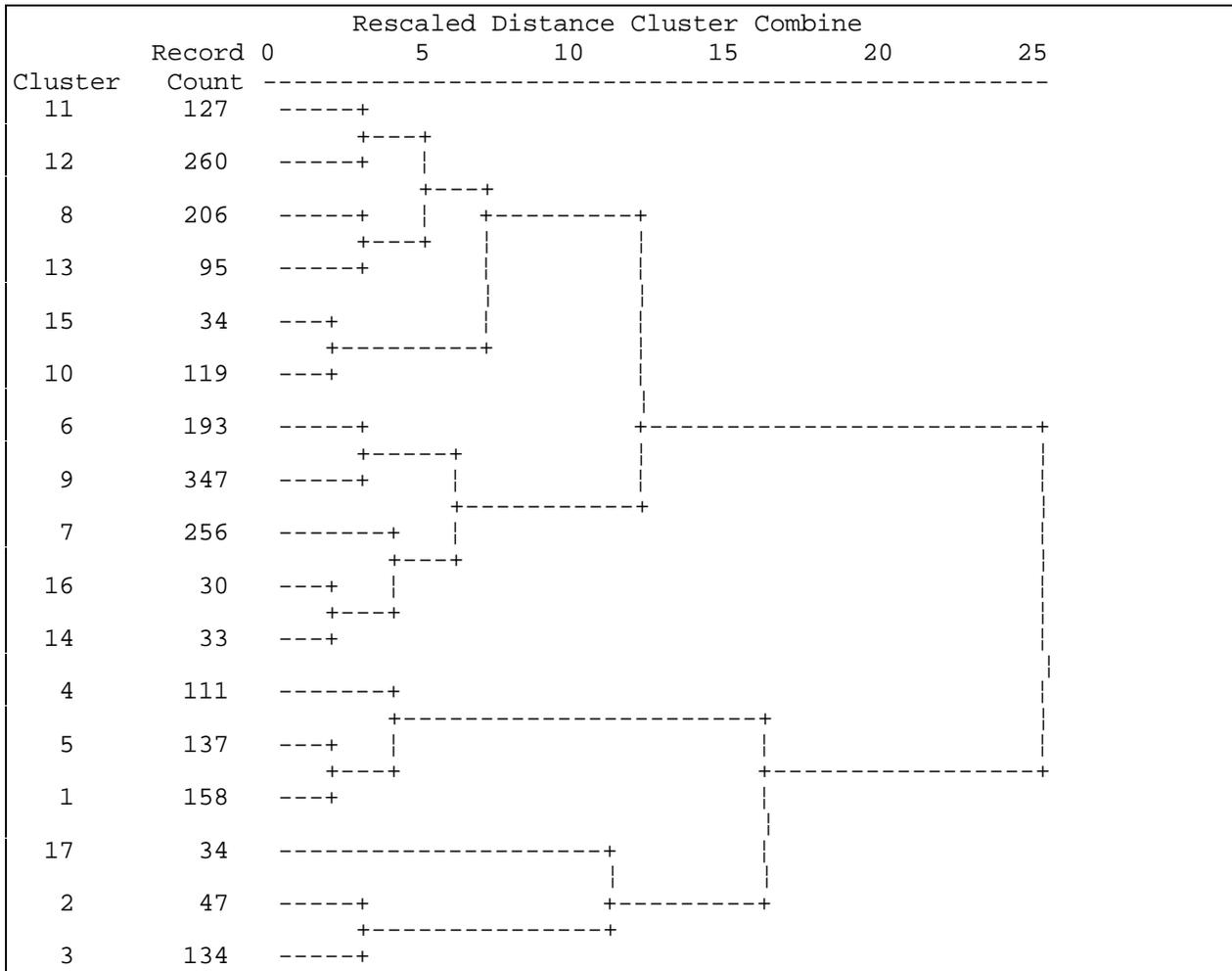


Table 2 provides summary information for the six most non-compliant clusters of the 17-cluster solution. Using the types of market segments included in a cluster and the compliance characteristics summarized below, subjective

descriptive names can be assigned to each of the clusters. Although this typically is done after cluster analysis, no attempt was made to label the clusters resulting from this analysis.

**Table 2. Average standardized factor scores for the six most non-compliant clusters.**

<i>Non-Compliant Cluster</i>	<b>Standardized Factor Scores</b>		
	<i>Payment</i>	<i>Accuracy</i>	<i>Size</i>
17	6.27	-0.91	1.08
16	2.64	-0.13	-0.54
15	-0.73	2.31	1.21
10	0.65	1.90	0.63
11	-0.58	1.35	-0.31
13	1.15	0.80	-0.25

*Cluster 17 — Large markets with accurate returns but totally unpaid balances.*

- Percent of returns unpaid at time of filing: 100%
- Percent of dollars unpaid at time of filing: 32%
- Average amount unpaid at time of filing: \$1,987
- Percent of returns accurately filed: 39%

*Cluster 16 — Small markets with unpaid balances.*

- Percent of returns unpaid at time of filing: 22%
- Percent of dollars unpaid at time of filing: 38%

*Cluster 15 — Large markets with inaccurate but fully paid returns.*

- Percent of returns unpaid at time of filing: 6%
- Percent of dollars unpaid at time of filing: 1%
- Average predicted tax change: \$3,410
- Percent of returns accurately filed: 18%

*Cluster 10 — Medium to large markets with inaccurate and unpaid returns.*

- Percent of returns unpaid at time of filing: 18%
- Average amount unpaid at time of filing: \$1,419
- Average predicted tax change: \$2,522
- Percent of returns accurately filed: 17%

*Cluster 11 — Small markets with inaccurate but paid returns.*

- Percent of returns unpaid at time of filing: 9%
- Percent of dollars unpaid at time of filing: 3%
- Average amount unpaid at time of filing: \$509
- Average predicted tax change: \$2,584
- Percent of returns accurately filed: 12%

*Cluster 13 — Small markets with inaccurate and unpaid returns.*

- Percent of returns unpaid at time of filing: 18%
- Percent of dollars unpaid at time of filing: 16%
- Average predicted tax change: \$2,138
- Percent of returns accurately filed: 11%

**Discussion**

The data reduction techniques presented in this paper were successful at consolidating a lot of data into useful information. The original data were represented by 11 measures for 71 different market segments across 33 geographic areas. In total, these were over 25,000 statistics. Factor and cluster analysis reduced this to 3 characteristics across 17 groups — 51 statistics. This becomes a much more interpretable set of data upon which to apply the

analysts' wisdom and insight. The compliance characteristics of the six non-compliant groups easily can be interpreted. With reference back to the original market segments, the analyst can formulate a comprehensive group comprised of the market segment intersections and overlaps.

Although the original categorization of the compliance measures seems to make sense, the factor analysis found a better allocation. Originally the totals for each attribute were spread across the categories — total population under market size, total taxes unpaid under timely payment, and total taxes reported along with total predicted tax increase under tax accuracy. The factor analysis concluded that these measures of totals more accurately represented a single characteristic — market size.

The factor analysis also demonstrated the limited usefulness of the VCL measure to identify non-compliance in this setting. The VCL measure had relatively little variability and did not associate or correlate with the other ten measures.

The factor analysis also demonstrated the indirect relationships of the measures. For example, from Table 1, it can be seen that average tax dollars unpaid has a moderate influence on the accuracy of a market segment: loading at 0.54. Conversely, the average predicted tax increase has relatively little influence on the timely payment of taxes. And lastly, while the percent of returns late has little in common with the characteristics identified (communality of 0.33), it does have a moderate influence on the payment characteristic, loading at 0.51.

The data reduction techniques used in this analysis can be applied to a wide variety of situations where the analyst is presented with large volumes of summary information that needs to be distilled into a smaller set of "key information".

## Acknowledgments

The authors would like to thank Michael Gregory of the North Central DORA and the entire Baseline National Compliance Measures Project Team. Additional thanks go to Teresa O'Hearn, our summer intern, for all her enthusiastic and helpful research assistance. A special thanks to Jeff Butler for his review and input.

---

## About the Author(s):

**Larry R. May** is an operations research analyst Team Leader in the Pennsylvania District Office of Research and Analysis. He received his B.S. degree in Business Administration (Accounting) in 1984 from Glassboro State College. He has been with the IRS since 1983.

**Anne Steuer** is a statistician in the North Central District Office of Research and Analysis. She received her Ph.D. degree in Natural Resource Economics and Policy in 1986 and her M.S. degree in Statistics in 1997 from the University of Minnesota. She has been with the IRS since 1995.

## Appendix A -- Description of Factor Analysis

Let the random variables  $X_1, X_2, \dots, X_{11}$  represent the eleven compliance measures and let the random vector  $X^T = [X_1, X_2, \dots, X_{11}]$  have covariance matrix  $\Sigma$  with eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{11} \geq 0$ . (Eigenvalues are scalars  $\lambda_1, \lambda_2, \dots, \lambda_k$  which are solutions to the polynomial equation  $|A - \lambda I| = 0$  where  $A$  is a  $k \times k$  matrix and  $I$  is the  $k \times k$  identity matrix. Details are available in any linear algebra text such as: Graybill, F.A. 1969. Introduction to Matrices with Applications in Statistics, Belmont, California: Wadsworth.) The  $i$ th principle component, or factor, is  $Y_i = e_i^T X$  where  $e_i$  is the  $i$ th eigenvector. (An eigenvector is a nonzero vector  $x$  such that for a  $k \times k$  matrix  $A$  and eigenvalue  $\lambda$ ,  $Ax = \lambda x$ . Since eigenvectors are, by convention, set to length one,  $e = \frac{x}{\sqrt{x^T x}}$  and is the eigenvector associated with the eigenvalue  $\lambda$ .)

Further the  $Var(Y_i) = e_i^T \Sigma e_i = \lambda_i$  and since  $Cov(Y_i, Y_j) = e_i^T \Sigma e_j = 0$ , the principle components, or factors, are uncorrelated.

The eleven compliance measures are varied from totals, averages, and percentages of both counts and dollars. Because of the wide range of measurement scales, the eleven measures were standardized prior to the factor analysis. The random vector  $X^T = [X_1, X_2, \dots, X_{11}]$  was transformed to  $Z = (V^{1/2})^{-1}(X - \mathbf{m})$  where  $V^{1/2}$  is an  $11 \times 11$  matrix with standard deviations on the diagonal and zeros on the off-diagonals and  $\mathbf{m}$  is a vector of means. This transformation is the familiar standardization formula  $z = \frac{x - \mathbf{m}}{\mathbf{s}}$  written in matrix notation. Here  $\mathbf{s}$  is the standard deviation and  $\mathbf{m}$  is the mean of the random variable  $x$ .

The random vector  $Z^T = [Z_1, Z_2, \dots, Z_{11}]$  has mean  $E(Z) = 0$  and  $Cov(Z) = (V^{1/2})^{-1} \Sigma (V^{1/2})^{-1} = \mathbf{r}$  where  $\mathbf{r}$  is the correlation matrix for the original random vector  $X^T = [X_1, X_2, \dots, X_{11}]$ . Thus the principle components or factors for the eleven standardized measures are derived from the correlation matrix for the unstandardized measures. These factors are generally not the same as the factors obtained using the untransformed variables. Thus standardization will yield different results. See Johnson, R.A. and Wichern, D.W. 1992. Applied Multivariate Statistical Analysis 3rd edition. Prentice-Hall, Inc.

## Appendix B -- Description of Cluster Analysis

Without a predetermined number of groups, a hierarchical approach was used to classify the cases into groups.<sup>32</sup> Specifically, Ward's Method clustering algorithm was used with the factor scores from the preceding factor analysis describing each market segment and Euclidean distance as the measure of similarity. Ward's method was selected since the distance measure is suitable for use with factor scores. Ward's method is biased towards clusters of similar sizes.<sup>33</sup>

The first step is to calculate the dissimilarity between two observations. This is calculated as the Euclidean distance between the two objects in P dimensional space where P is the number of characteristics being considered.

Although Euclidean distance is a common measure of similarity in cluster analysis, it has two weaknesses: the calculation is sensitive to magnitudes of scale and correlation between the variables<sup>34</sup>. In this baseline study, both of these concerns are alleviated as a result of the factor analysis. This is because, as previously discussed, factor scores are standardized; hence no scale problems. Furthermore, factor scores

<sup>32</sup> Cluster analysis techniques can be broken down into two general categories; hierarchical procedures and nonhierarchical procedures. In hierarchical procedures the technique moves up through the grouping of individual cases into clusters until all cases are in one group (agglomerative methods) or down through all cases in one group to all cases in separate groups (divisive methods). The decision at each stage of which clusters to combine or split is based on the similarity measure. In nonhierarchical procedures a specified number of clusters or maximum distance is established. Observations that meet the algorithm's criteria are assigned to the closest cluster. A critical step in the nonhierarchical procedures is the specification of cluster centers. See Joseph F. Hair, Jr. et al., *Multivariate Data Analysis: with Readings* -- 4th ed. (Upper Saddle River, New Jersey: Prentice-Hall) for a complete discussion of the common clustering techniques.

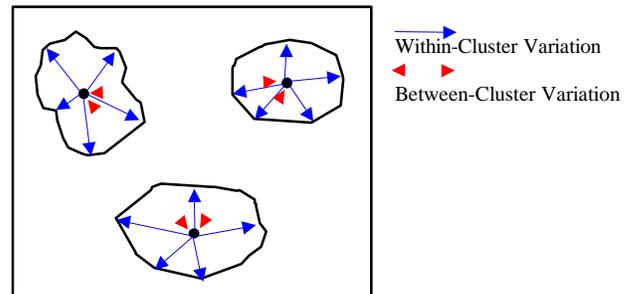
<sup>33</sup> Joseph F. Hair, Jr. et al., *Multivariate Data Analysis: with Readings*, 4th ed., (Upper Saddle River, New Jersey: Prentice-Hall), p. 440.

<sup>34</sup> Brian S. Everitt, *Cluster Analysis*, 3rd ed., (New York: Halsted Press), pp. 46-47.

are jointly uncorrelated; hence no collinearity problem.

After the dissimilarity is quantified, the data are analyzed to group similar cases. In Ward's Method this is done by evaluating the within-cluster variation versus the between-cluster variation. Figure B1 demonstrates this concept<sup>35</sup>

Figure B1: Within vs. Between Cluster Variation



$$d_{ij} = \sqrt{\sum_{k=1}^p [(X_{jk} - X_{ik})^2]}$$

In hierarchical methods of cluster analysis the grouping of data is exhaustive. Clustering continues until all the data is grouped into one cluster. For example the above diagram shows three groups. At the next stage of the analysis two of these groups would be joined leaving two clusters, and in the final stage those two clusters would be joined into one collective group. While individual cases may be too much data to interpret, a single cluster provides little information as well. The optimum number of clusters is somewhere in between. Determining the optimum number of clusters is a matter of considerable debate<sup>36</sup>. This remains a subject of deliberation because the context of the clustering application and the structure of the data have at least as

<sup>35</sup> Adopted from Joseph F. Hair, Jr. et al., *Multivariate Data Analysis: with Readings*, 4th ed., (Upper Saddle River, New Jersey: Prentice-Hall), p. 437.

<sup>36</sup> See Milligan, G. & Cooper, M. 1985, "An Examination of Procedures for Determining the Number of Clusters in a Data Set," *Psychometrika*, Vol. 50, No. 2, pp. 159-179. and Dubes, R. & Jain, A. (1979), "Validity Studies in Clustering Methodologies," *Pattern Recognition*, Vol. 11, pp 235-254.

much influence on the decision as mathematical theory and the technique applied.

In this analysis, traditional cluster output of dendrograms and agglomeration schedules were reviewed along with summary statistics for each group at each stage of the clustering process. Since the objective of this analysis was to identify groups of non-compliant market segments at the National level, three subjective operational criteria also were considered in determining the final number of clusters:

Optimization Criteria

1. A non-compliant cluster was defined as having a median factor score for either the Payment or Accuracy characteristic in excess of 0.8.
2. The collective number of market segments contained in the non-compliant clusters was approximately 10.
3. A market segment was considered part of a cluster if at least half of the districts appeared in that cluster.

The cluster solution that minimized the overall number of clusters while meeting the three criteria above was deemed the optimum solution.

# Review of the IRS's Individual Return Electronic Filing and Related Research

By Javier Framinan

The Internal Revenue Service has profiled and studied individual income tax return filers in an effort to learn how to improve and market its electronic filing products. This article reviews and summarizes that research, examining who files electronically, why they do, why they do not, the costs of electronic filing, and the results of various electronic filing marketing initiatives. Overall, the various research efforts show electronic filers are young, have lower income, and have simple tax situations compared to the general individual filer population. In 1998, 20 percent of individual income tax returns were filed electronically. The biggest motivation for filing electronically is a fast refund. Not surprisingly, over 30 percent of taxpayers with refunds in the \$1,000 to \$3,000 range elect to file electronically. Over 40 percent of taxpayers receiving the Earned Income Tax Credit file electronically. The IRS wants to maintain a current knowledge base of the individual taxpayer market segment in order to hone its electronic filing marketing strategy. Also, as the IRS shifts some of its attention to the electronic filing of business returns, it will need to conduct similar research in the profiling and study of business entities for strategic planning purposes.

## Introduction

The Internal Revenue Service recognizes that manual processing of paper-based interaction with the public is a resource-intensive activity, prone to errors. (The "public" includes individual and business taxpayers, taxpayer representatives, tax practitioners, and other government entities.) Data entry error rates for paper individual income tax returns are approximately 20 percent, compared to about 2 percent for electronically filed returns.<sup>37</sup> The IRS also recognizes taxpayer information (stored in IRS databases) will improve with electronic commerce and communications. The IRS captures 100 percent of the information recorded on electronically filed returns, compared with approximately 40 percent from paper returns.<sup>38</sup> More information and easier access to it, in turn, improves the IRS's ability to serve the public. In addition, from the public's perspective, electronic exchange offers easier ways to file returns and pay tax liabilities, confirms the IRS's receipt of returns, speeds up refunds, and enables electronic retrieval of

forms, publications and other information from the IRS (e.g., from the IRS Internet homepage).

This paper gives a brief historical review of the IRS's electronic filing activities with individual income taxpayers, but primarily serves to review the research conducted to date. In general, the goal of the research related to electronic filing, or *e-file*, has been to understand its market segments and increase participation. The discussion is limited to the electronic filing of individual income tax returns, as this is where the IRS has concentrated its efforts. In fact, this review of the literature confirms the limited research done to date on the electronic filing of business returns, and points to a clear need for more IRS study in that area.

## Background

In 1998, 20 percent of individual income tax returns were filed electronically. However, in that same year the Congress stated its intent that by 2007 the IRS should conduct 80 percent of its interactions with the public electronically.<sup>39</sup> This will require aggressive development of the IRS's electronic commerce. The Office of the Assistant Commissioner (Electronic Tax Administration), or ETA, is the lead IRS organization

<sup>37</sup> Internal Revenue Service, Electronic Tax Administration. (1997). *Critical Issues for Development of an Electronic Tax Administration Strategy: A Plan for Moving Federal Tax Administration into the Information Age*. Washington, DC. & General Accounting Office. (1993). *Opportunities to Increase the Use of Electronic Filing*. Washington, DC.

<sup>38</sup> General Accounting Office. (1995). *Electronic Filing Falling Short of Expectations*. Washington, DC.

<sup>39</sup> The IRS Restructuring and Reform Act of 1998.

in attempting to meet this challenging 80 percent goal. In December 1998, it submitted to the Congress its first installment on its strategic plan to revolutionize how taxpayers transact and communicate with the IRS.<sup>40</sup>

The ETA's strategic planning depends on an understanding of its customers -- who uses the currently available electronic commerce products and why, and who does not and why. Data-driven research helps the ETA understand the technological, legal, and financial barriers facing tax-related electronic commerce. It uses research results to identify programs that give IRS customers incentive to shift from traditional (i.e., paper-based) interaction to electronic.

### Individual Return Electronic Filing

The IRS started electronic filing of individual income tax returns as a test in 1986, and expanded it nationwide in 1990. Individual electronic filing is one of the IRS's most developed forms of electronic commerce. Consequently, most of the IRS's electronic filing research relates to the individual income tax returns. Individual e-file can be divided into two distinct categories: "standard e-file" and TeleFile. Until 1994, standard e-file included only electronic returns prepared and filed for the taxpayer by an authorized professional tax practitioner, or returns prepared by taxpayers themselves and taken to an approved transmitter who transformed the return information into the necessary electronic format and transmitted the return electronically. (This "standard e-file" also was known as "ELF," or Electronically Filed.) Presently, it also includes On-Line filing, comprised of returns prepared by taxpayers that use tax preparation software and transmit on-line through an authorized electronic return filer. TeleFile returns include all those transmitted over the telephone using touch-tone technology. Table 1 presents historical annual volumes of individual e-file, as well as its overall market penetration rates. Also, more detailed descriptions of these categories are provided in the gray box.

<sup>40</sup> Internal Revenue Service, Electronic Tax Administration. (1998). *A Strategy for Growth*. Washington, DC.

#### Individual Return e-file Categories and Sub-categories

##### *Electronic Return Originator/Practitioner e-filing*

In 1986, before the proliferation of personal computers and modems, individual taxpayers could file electronically only if they went to a professional tax practitioner (also referred to as a preparer). The IRS coordinated with electronic return originators (EROs) and the preparer community, and set technical and procedural standards for electronic information exchange. Although taxpayers incurred additional fees filing electronically, faster refunds provided incentive for many to e-file. By 1992, after only two years of nationwide implementation, almost 11 million taxpayers were filing electronically through a preparer. The attraction of faster refunds gave rise to the refund anticipation loan (RAL) market. For an extra fee (in addition to the return preparation and transmission charges), electronic filers could secure a RAL, where in coordination with the preparer, a bank advances the anticipated tax refund amount. In essence, this gives the taxpayer an instant refund upon tax return transmittal.

##### *On-Line Filing*

**On-Line filing has grown with the popularity of tax preparation software. In 1996, in its second year of existence, 158,000 taxpayers participated in the On-Line filing program. By 1998, 942,000 were filing using this method. The volume of On-Line filers is projected to grow to over 2 million in 1999.<sup>41</sup>**

To file on-line, the taxpayer must have a computer, modem, and tax preparation software from a certified private vendor. On-Line filing also requires use of an IRS-accepted on-line service company or transmitter. After completing a tax form electronically, the on-line filer pays a fee directly to a return transmitter company (or indirectly to the transmitter, through the purchase of the tax preparation software) that translates the return information into an IRS readable format. On-Line filing basically has the same incentives and restrictions as practitioner electronic filing -- i.e., faster refunds, higher accuracy, IRS confirmation of receipt, but extra cost.

##### *Electronically Filed, but not Prepared*

A small number of taxpayers take a hard copy of their returns to a practitioner or other transmitter to submit it electronically to the IRS. During filing year 1997, 2.1 percent of all individual tax returns were filed this way.<sup>42</sup> Again, the taxpayer motivation is an accelerated refund.

##### *TeleFile*

In 1996, the IRS offered TeleFile nationwide. TeleFile employs touch-tone telephone technology to transmit returns, using IRS-issued customer service numbers for authentication. The IRS has limited TeleFile's availability to filers of simple returns (i.e., Form 1040EZ), as research shows taxpayers' unwillingness to key enter information for longer returns on a touch-tone pad. This filing option caters to taxpayers unwilling to use a preparer and/or unwilling to pay transmission fees. Every year since 1996, the IRS has mailed TeleFile tax packages to approximately 25 million taxpayers identified as eligible. In 1998, almost 6.0 million taxpayers filed this way, making up 24 percent of the individual return e-file market.

<sup>41</sup> Internal Revenue Service, Research and Statistics of Income. (1998). *Calendar Year Projections of Individual Returns by Major Processing Categories Selected Years and Areas Fall 1998 Update*. Washington, DC.

<sup>42</sup> Internal Revenue Service, Southwest District Office Research and Analysis. (1997). *A National Profile of IRS e-file Users in 1997*. Phoenix, AZ.

**Table 1. Electronically Filed Individual Returns: 1986 - 1998**

Filing Year	Standard e-file	TeleFile	On-Line Filing*	Total e-file	e-file as a Percent of All Individual Returns
1986	24,814			24,814	0.02%
1987	77,612			77,612	0.08%
1988	583,462			583,462	0.55%
1989	1,160,516			1,160,516	1.06%
1990	4,204,188†			4,204,188	3.74%
1991	7,567,116	Did Not Exist		7,567,116	6.65%
1992	10,919,281	125,981		11,045,262	9.63%
1993	12,333,750	148,585	Did Not Exist	12,482,335	10.97%
1994	13,502,055	518,693	285	14,020,748	12.23%
1995	11,126,885‡	680,010	1,372†	11,806,895	10.17%
1996	12,128,969	2,839,437†	158,284	14,968,406	12.65%
1997	14,449,712	4,685,959	366,727	19,135,671	15.90%
1998	18,625,689	5,954,564	942,176	24,580,253	20.06%

\* On-Line Filing volumes are a subset of "Standard e-file."  
 † Indicates first year of nationwide implementation.  
 ‡ The 1995 drop in standard e-file volume was due to the IRS Revenue Protection Strategy, instituted to combat refund fraud associated with electronic filing.

**Who Files Individual Returns Electronically**

*Initial Research*

The IRS understood relatively little about its individual electronic filing market when it started the program in 1986. The Service depended primarily on the tax practitioner community to promote its growth. However, by the 1990s, the IRS recognized the need for research, not only for marketing purposes, but also for resource allocation planning. (With taxpayers switching from paper to electronic filing, resources formerly committed to the manual processing of returns need to be reallocated.)

In 1991, the IRS Research Bulletin featured an article by Bryan Musselman that profiled individual return electronic filers. The article described electronic filers from the 1990 and 1991 filing seasons in terms of age, education, income, geographic location, and form type. Since the IRS lacked internal data related to electronic filers, the article relied heavily on survey data gathered by the Roper Organization. Despite the reliance on external data, much of the article's general profile and

conclusions remain valid today. The research found the typical individual electronic filer files a relatively simple return, is young, has a lower income, most likely lives in the southeast, and is motivated by a quick refund. Barriers to electronic filing are awareness and cost.

*A Lack of Information Leads to e-file Research Database*

The absence of an IRS database for profiling taxpayers eligible to file electronically hindered efforts to market e-filing during the early to mid 1990s. IRS data to corroborate the 1991 Roper survey results were scant, and any annual changes in the e-filer profile were difficult to determine. These deficiencies did not go unnoticed. The General Accounting Office's *Tax Administration, Electronic Filing Falling Short of Expectations* (1995) concluded the lack of adequate data and inability to perform cost/benefit analyses hampers IRS decision-makers, and contributes to their lack of strategic focus. A year later, the ELF Profiling Project Team's *Profile Report: Current and Potential Market Segments for Electronic Filing* (1996) warned "further research and analysis of the electronic filing program will be hindered by the lack of a timely database."<sup>43</sup> The report recommended development of a database to enhance the profiling of electronically filed returns. By 1997, the IRS Southwest District Office Research and Analysis (DORA) was tasked with coordinating research in the "alternative ways of filing" area (i.e., electronic filing). However, in its *National Profile of IRS e-file Users in 1997* (1997), the DORA admitted to an as yet inadequate understanding of the individual e-file market segments. The researchers cited the need to build an individual e-file database to enable the study of the market segment.

In 1997, the Southwest DORA took a step toward filling the data void, by constructing the "national e-file research database," which contained tax year (TY) 1996 individual income tax return data on characteristics such as taxpayer location, age, income, taxes, deductions and exemptions. The database was "intended as a prototype for an annual

<sup>43</sup> Internal Revenue Service, ELF Profiling Project Team. (1996). *Profile Report: Current and Potential Markets for Electronic Filing*. Washington, DC.

construction of similar databases.”<sup>44</sup> The report describes the database development and sources. It also considers research questions posed by the ETA, and profiles the practitioner, On-Line, and TeleFile market segments, as well as the paper return filers for comparison purposes. More recently, the Southwest DORA updated the national e-file research database, and renamed it the “ETA Market Research Database” (see *On-Line Filers: ETA Market Research for 1998* (1999)).

*Profiles and Other Facts Regarding Individual Taxpayers*

The IRS has conducted several studies, and contracted private vendors to conduct several more, to understand its e-file markets. Much of this research has been of a profiling nature – i.e., it describes the demographic characteristics of taxpayers who e-file, as well as eligible taxpayers who do not. The findings vary little, even across time. Following are various e-filer demographics as reported in two major profiling reports: *Profile Report: Current and Potential Market Segments for Electronic Filing* (1996) and *A National Profile of IRS e-file Users in 1997* (1997).

Profile Report: Current and Potential Market Segments for Electronic Filing

In 1995, prior to the establishment of the ETA, the IRS's electronic filing executive requested the IRS Director, Compliance Research profile electronic filing market segments. Despite the absence of a database specifically designed to study the e-file market, the National Office Research and Analysis (NORA) and the ELF Profiling Project Team (consisting of IRS National Office and District Office representatives) under Compliance Research conducted a national profile of taxpayers filing in 1994 and 1995. Meanwhile, the IRS's District Offices Research and Analysis (DORAs) produced similar local profiles. The *Profile Report: Current and Potential Market Segments for Electronic Filing* (1996) reported findings from both efforts, predominantly on returns filed in 1994. This research profiled all taxpayers eligible to file electronically, including those that did and those that did not, using TY 1993

<sup>44</sup> Internal Revenue Service, Southwest District Office Research and Analysis. (1997). *A National Profile of IRS e-file Users in 1997*. Phoenix, AZ.

sample data from the interim Compliance Research Information System (CRIS) file and data from the Automated Wage Information File (Autowif). The analysis excluded TeleFile, since it still was not available nationwide at the time.

The resulting report essentially confirmed Musselman's 1991 research in profiling the electronic taxpayer. Electronic filers of individual income tax returns are predominantly younger, lower income, simpler return, and motivated by a faster refund. The report cited 98.4 percent of TY 1993 individual return filers were eligible to file electronically; however, only 12.2 percent actually did.

Tables 2 through 5 present selected statistics from the *Profile Report: Current and Potential Market Segments for Electronic Filing* (1996). They reflect returns filed in 1994 (for TY 1993) and show participation among *eligible* taxpayers by category. In Table 2, for instance, of all tax returns eligible for e-file with adjusted gross income (AGI) less than \$13,000, 14.3 percent actually were e-filed.

**Table 2. e-file Penetration by Adjusted Gross Income**

Adjusted Gross Income	Participation Rate of Eligibles
< \$13,000	14.3%
\$13,001 - \$26,000	16.3%
\$26,001 - \$39,000	10.8%
\$39,001 - \$52,000	9.2%
> \$52,001	5.1%

**Table 3. e-file Penetration by Age**

Age	Participation Rate of Eligibles
< 16	0.6%
16 – 20	8.1%
21 - 24	18.8%
25 – 44	17.2%
45 – 64	7.9%
> 64	2.2%

As Tables 2 and 3 show, e-file participation drops dramatically among taxpayers with AGI above \$26,000 and older than 44 years of age.

Table 4 presents information contained in the report regarding return complexity and its relationship to e-file participation. In general, e-filers are characterized by simpler returns. Stated differently, e-filing does not attract taxpayers with more complex income tax situations.

**Table 4. e-file Penetration by Return Complexity**

Return Schedules	Participation Rate of Eligibles
No Schedules	15.3%
Schedule A	7.2%
Schedule C	5.7%
Schedule F	3.9%
More than One Schedule	4.5%

The report also highlights various other characteristics regarding e-file, some of which are contained in Table 5. Not surprisingly, only 0.7 percent of eligible TY 1993 balance due returns were filed electronically. Also not surprisingly, of all taxpayers receiving a refund of less than \$300, only 2.8 percent filed electronically. The relatively high costs of electronic filing dissuaded these taxpayers, as the net refund amounts after payment of e-filing fees would be small. Conversely, of those taxpayers receiving a \$1,001 to \$3,000 refund, 31.4 percent apparently desired a fast refund and could justify the fees. The profile of the earned income tax credit (EITC) recipients tells the same story. (The earned income tax credit is a Federal government credit provided to lower income individuals. It is not related to income tax, but is deducted from income tax liabilities and disbursed through income tax refunds.) Over 42 percent of the taxpayers receiving the EITC e-filed to get their money faster. The larger check from the government apparently offsets the costs associated with e-filing.

**Table 5. Miscellaneous e-file Penetration Characteristics**

Characteristic	Participation Rate of Eligibles
Refund Return	17.2%
Balance Due Return	0.7%
Paid Prepared Return	18.0%
Non-Paid Prepared Return	6.6%
Refund < \$300	2.8%
Refund \$300 - \$1,000	16.2%
Refund \$1,001 - \$3,000	31.4%
Refund > \$3,000	16.6%
Form 1040 – Type Return	12.3%
Form 1040A – Type Return	19.0%
Form 1040 EZ – Type Return	6.6%
Single Filing Status Return	7.0%
Joint Filing Status Return	9.4%
Head of Household Filing Status Return	38.3%
Return with Earned Income Tax Credit	42.1%
Return without Earned Income Tax Credit	7.6%

A National Profile of IRS e-file Users in 1997

The IRS's Southwest District Office Research and Analysis conducted further individual e-file market profiling on TY 1995 and 1996 returns and produced *A National Profile of IRS e-file Users in 1997* (1997). The report confirmed most of the e-filer profile information presented in earlier research, but contains more in terms of demographic detail and distinction between "standard e-file," On-Line filing, and TeleFile.

According to the 1997 report, TeleFilers are young, single, have low income, and have low refund amounts (compared to the average for all individual returns filed). Non On-Line standard e-filers (i.e., originating from EROs/practitioners) are a little older, have higher incomes and refunds, and consist of more EITC recipients. On-Line filers have even higher income, higher refunds, and more complex returns.

Following are some of the 1997 report's general findings, by demographic characteristic.

*Age:* E-filers are predominantly younger. The largest age group of TeleFilers was concentrated in the 18 to 29 year old range, while the majority of standard e-filers was in the 30 to 44 year old range. Conversely, paper return filers are more evenly distributed by age.

*Income:* E-filers' average income is lower than that of the entire taxpayer population's. More TeleFilers and non On-Line filers fell in the less than \$15,000 adjusted gross income category than in any other AGI group defined in the study. However, On-Line filers have higher AGIs. Forty-five percent of On-Line filers had an AGI of \$50,000 or more.

*Refunds:* Standard e-file returns (not including On-Line) had an average refund of over \$1,827 in 1997, while TeleFile returns averaged \$400. The average refund for all individual returns filed that year was \$627. The researchers attribute the high refunds for non On-Line standard e-file returns, at least in part, to the earned income tax credit from the government. The e-file returns originating from tax preparers were characterized by lower income, yet higher refunds due to the EITC payments. Almost 49 percent of the non On-Line standard e-file returns claimed EITC. The average EITC payment for all non On-Line standard e-file returns (including returns not claiming EITC) was \$880, compared to On-Line filers' (the next closest group) \$214. The combination of low income and high refund amount apparently creates the demand for quick refunds, regardless of the extra preparation and filing expense. This phenomenon holds true for the higher income market segment characterizing the On-Line filers, as well. The *On-Line Filing Program: Focus Group Report (1997)* by the IRS Office of Opinion Research concluded some taxpayers more readily pay the extra cost of e-filing if they expect a large refund check.

*Professionally Prepared:* Almost 49 percent of all individual income tax returns filed in 1997 were prepared professionally. Professionals prepared 81 percent of all electronically filed returns, compared to 42 percent of all paper returns.

*Return Complexity:* On-Line filed returns were

the most complex of all returns (paper and electronic combined). The researchers used the number of schedules attached as an indicator of complexity. They found 80.2 percent of On-Line returns had at least one schedule attached. Further IRS success with On-Line filing would address concerns noted by the U.S. General Accounting Office (GAO). In 1995 the GAO reported only about 20 percent of the e-file returns in 1994 were Form 1040-type -- i.e., more complex -- even though 59 percent of all individual returns filed in 1994 were Forms 1040. The GAO report suggests IRS had focused on an e-file return volume goal, rather than an operating cost reduction one. It recommends the marketing strategy consider targeting taxpayers with more complex returns that are more expensive for the IRS to process manually.

*Repeat Rates:* In general, taxpayers remain loyal to their filing method from year to year. Paper filers had the highest repeat rate, at 90 percent, while On-Line filers showed the lowest, with 55 percent. (*On-Line Filers: ETA Market Research for 1998 (1999)* subsequently profiled On-Line filers for TY 1996 and 1997. It found overall the repeater rate rose to 59 percent for the period from TY 1996 to 1997 filings.) For non-On-Line standard e-file returns the repeater rate is 74 percent. For TeleFile, the overall repeater rate is 56 percent. Many TeleFile users, however, can not repeat the following year due to TeleFile restrictions on income level, complexity of returns, change of address, age, and dependents.

*Regional Differences:* There are some regional differences with respect to electronic filing. The profile confirmed earlier studies that the southeast experiences a higher participation rate (24 percent), while the west has a lower rate (13 percent). Returns e-filed from the IRS's Western Region were more complex, and had lower refunds. E-file returns from the Southeast Region had the highest refunds, on average, and had a higher rate claiming the EITC. The study defined the TeleFile participation rate as TeleFile returns filed divided by TeleFile packages mailed. The Northeast Region experienced the highest TeleFile participation rate at almost 20 percent, compared to Western Region's 15 percent. (However, an exceptionally aggressive state-level telefile program in Massachusetts has a unique effect on the Federal TeleFile participation rate in that state, boosting it to almost twice the national

average. The author suggests this situation likely contributes to the study's findings relative to the Northeast Region.)

**Other Profiling Research**

The Pacific Consulting Group profile of TY 1995 On-Line filers, *Minimizing Taxpayer Burden and Reducing Costs at the IRS: Analysis of Customer Experience with On-Line Filing* (1996), identified On-Line filers as a very specialized group. The survey research found the On-Line filer to be computer literate, highly educated, and with a high income. Unlike the market of lower income e-filers that uses a practitioner, the On-Line filer market is more savvy and will require more incentive (in the form of lower perceived costs) to e-file.

A Bonney & Co. survey of TY 1995 TeleFilers (1996) found these filers to be young and somewhat educated, and identified a substantial untapped market of less educated eligible taxpayers that lack confidence using a computerized data entry system. The report suggests the IRS direct TeleFile marketing efforts at these non-participants to overcome their reluctance.

**Cost/Benefit of e-file**

All of the IRS's efforts in the electronic commerce arena are predicated on the assumption such activity will save resources. However, little (and incomplete) cost-benefit analyses have been conducted to show definitively electronic filing, and electronic commerce in general, saves IRS resources in the short term. One of the reasons for the lack of research in this area is the complexity of the task. Estimation of upstream and downstream costs and benefits (e.g., those related to facilities, equipment, storage, archiving, subsequent adjustment activity, training, audits and other compliance-related activities) requires data not readily available to the IRS, or techniques as yet undeveloped. Nevertheless, the GAO (1995) reported the following per return processing costs for the 1993 filing season.

1993 Individual return processing costs (per GAO):

- Form 1040 - \$4.53
- Form 1040A - \$3.95
- Form 1040EZ - \$3.36
- e-file (excluding TeleFile) - \$3.08

These cost figures do not consider the up- and downstream cost, but just those associated with return processing – i.e., the opening and sorting of mail, coding, editing, data entry, validity checks, and error correction at the IRS Service Centers and Processing Centers. Examples of upstream costs include design, printing, and distribution of tax forms and instructions; examples of downstream costs include archiving and retrieval of tax returns, and compliance activities.

Steuer and Benson (1996) attempted to rank e-file market segments based on cost savings. They used the GAO's per return cost data (along with other processing year 1994 data on individual taxpayers in the state of Minnesota) in a "tree-structured analysis" that used SPSS CHAID, or Chi-squared Automatic Interaction Detector. They defined e-filer market segments along the following taxpayer/return characteristics: adjusted gross income, age, refund/balance due, tax credits (including EITC), paid preparer/self prepared, and geographic location. The analysis identified some inefficiencies in the IRS's e-file marketing. The research suggests the ETA use ranking techniques such as those based on CHAID to prioritize market segments and direct its e-file marketing strategies with a cost-benefit orientation.

More recent work by the Office of Cost Analysis under the IRS's Chief Financial Officer estimated the following per return processing costs for fiscal year (FY) 1996.<sup>45</sup>

FY 1996 Individual return processing costs (per CFO):

- Form 1040 - \$4.44
- Form 1040A - \$3.58
- Form 1040EZ - \$3.54
- Form 1040PC - \$3.44
- e-file (excluding TeleFile) - \$4.73
- TeleFile - \$3.88

These figures also exclude up- and downstream costs; but they show e-filing is relatively more expensive than paper filing alternatives. Further preliminary analysis by the CFO suggests inclusion of the up- and downstream costs would

<sup>45</sup> Internal Revenue Service, Office of Cost Analysis. (1998). *FY 1996 Cost of Submission Processing in the Service Centers*. Washington, DC.

not change this conclusion. However, the IRS, and the ETA in particular, argue e-file's cost effectiveness will be realized with economies of scale, when e-filing becomes the predominant form of filing and the large sunken costs associated with the IRS's paper-based system are eliminated.

Also, it is important to note the processing costs for paper returns in the studies described above are based on a system that captures only about 40 percent of the information on returns. In contrast, the e-file cost figures reflect 100 percent data collection. A more balanced comparison requiring 100 percent data collection for paper processing would increase significantly the corresponding cost figures for paper returns cited above.

Still, more complete cost/benefit analysis is necessary to enable IRS management, as well as the Congress, to make informed business decisions regarding e-file expansion.

**After the IRS Knows *Who* e-files, It Must Study *Why***

While profiling efforts enable IRS researchers to describe taxpayers that electronically file and those that do not, they can not fully explain taxpayers' motivations. Why does a taxpayer choose to file electronically? More importantly, why does a taxpayer not file electronically? The following discussion addresses each e-file category (i.e., ERO/practitioner e-file, On-Line, and TeleFile) separately.

*Why Do Taxpayers e-file?*

**ERO/Practitioner Standard e-filing**

After describing electronic filers as young, lower income, simpler return taxpayers, Musselman (1991) speculated various reasons for their participation. "First, most electronic transmitters also offer a 'refund anticipation loan' . . . Although this option generally costs approximately \$60 . . ., it requires no upfront cash, and thus may be the only way many people can afford tax preparation services. Second, the temptation to have one's money in

a matter of days instead of weeks may be a powerful incentive for many . . ."<sup>46</sup>

In its report to the U.S. Senate Finance Committee, the GAO (1993) cited the same latter motivation behind e-file participation: quicker refunds. This study involved GAO visits and interviews at four regional IRS offices, eight district offices, and three service centers, as well as a survey of over 1,000 preparers and transmitters that participated in electronic filing and 1,000 that did not.

**On-Line Filing**

The IRS's Strategic Planning Division's *Focus Group Report: On-Line Filing Program* (1997) cited quicker refunds and less paper to keep as On-Line filing selling points. The focus group participants (made up of taxpayers who had access to a computer and modem but did not e-file on-line) also perceived IRS acknowledgment that a return is received as an advantage of On-

Line filing. An earlier study by the Klemm Analysis Group (1996) and Pacific Consulting Group (1996) had similar findings. Their survey found accuracy of the filed return most influenced the decision to use the On-Line program. Also, they found convenience and speed of refund as the top "filing image items." *On-Line Filers: ETA Market Research for 1998* (1999) also found repeat On-Line filers are attracted by conveniences such as the ability to direct deposit their refunds.

**TeleFile**

The Bonney & Co. focus group (1996) and survey (1996) research cite ease of use, faster refunds, and no need for a paid preparer as taxpayers' incentives for using TeleFile. Earlier profiling work from the Southwest DORA supports this conclusion.

<sup>46</sup> Internal Revenue Service, Planning and Research. (1991). *Electronic Filing – Who's Participating and Who Isn't*. Washington, DC.

*Why Taxpayers Do Not e-file -- Barriers to Electronic Filing*

ERO/Practitioner Standard e-filing

Even before the IRS had a complete picture of *who* files electronically (i.e., before it had a comprehensive market profile), it began to study *why* taxpayer behave the way they do, and specifically the barriers to participation growth. In 1991, Musselman cited Roper survey data suggesting unawareness was electronic filing's biggest barrier. This seems logical in the early years of electronic filing. The Roper survey found 34 percent of individual taxpayers were not aware of the possibility to file electronically. The next largest barrier to the return filer was cost, cited by 14 percent of taxpayers.

As electronic filing became better known, cost replaced unawareness as the major impediment to participation. The GAO (1993) cited cost as e-file's main deterrent. Almost three years later, in its report to the U.S. Senate Government Affairs Committee, the GAO (1995) concluded the same. Taxpayers had to pay \$15 - \$40 to file electronically through a preparer or electronic filing transmitter. The GAO concluded e-filing appeals primarily to taxpayers most in need of a quick refund – those that disregard cost considerations.

Tax practitioners have had their own barriers to the electronic filing business. Nelco Inc.'s *Why Tax Preparers Do Not Offer Electronic Filing* (1994) explored the reasons for preparers' reluctance to e-file. Its survey research concluded mainly two interrelated factors affect participation: client demand and cost. The research recognizes the many dimensions of practitioners' cost – costs related to the purchase of software and hardware, transmission, training, etc. The preparers eventually must pass these costs onto the client. Higher fees, in turn, put the preparer at a competitive disadvantage if there is little client demand for e-file. Unwillingness to learn the new system and lack of confidence in it were two other reasons cited, though less frequently, by the survey respondents.

In September 1994, the Chairman of the American Institute of Certified Public Accountants (AICPA) provided GAO with the

following corroborating information on practitioners' barriers to e-filing:<sup>47</sup>

- e-filing does not fit into their "office routine,"
- clients do not perceive any additional benefit to offset the additional cost,
- e-filing requires additional input, transmitting, and monitoring time,
- e-file is not yet truly paperless – preparers particularly have a problem with the signature requirement. (Until filing year 1999, all such electronic filing still required the practitioner to prepare, sign and *mail* a paper signature document to authenticate each return, as well as attach Form W-2 earning statements and other documents requiring signatures.)

On-Line Filing

The proliferation of personal computers in the 1990s has created a huge potential for the On-Line electronic filing market. However, when the IRS made On-Line e-filing available nationwide in 1995, there was little promotion aimed at On-Line filers either from the software vendors or the IRS. Some tax preparation software vendors offered electronic filing at no extra cost (e.g., "first one free" offers), but this feature was not advertised by the vendors as a major selling point.

The 1996 On-Line filer customer satisfaction survey research by the Pacific Consulting Group and the Klemm Analysis Group found satisfaction among current users generally high, but also found high retransmission rates and low customer support ratings discourage repeat use or initial entrance into the program.

To devise an effective marketing plan to attract On-Line filers, the IRS conducted focus groups with eligible taxpayers that did not file On-Line to determine barriers. The Strategic Planning Division's *Focus Group Report: On-Line Filing Program* (1997) reported taxpayer awareness as the biggest issue. Only half the focus group participants knew the existence of On-Line filing, and then not much beyond that. General lack of knowledge – how the program works, its requirements, and where to get more information – manifested itself into fear and anxiety. Participants expressed concern

<sup>47</sup> General Accounting Office. (1995). *Electronic Filing Falling Short of Expectations*. Washington, DC.

regarding data security, technical problems with transmission, lack of paper documentation, as well as software and transmission costs. Some cited their techno-phobia associated with the use of computers on so complicated a task as tax return preparation. Also, some thought since electronically transmitted information is easier to access, it would be easier for the IRS to conduct random audits on such returns, and thus considered electronic filing disadvantageous.

### TeleFile

The Bonney & Co. focus group (1996) and survey (1996) research discovered similar fears among potential TeleFile users. Anxiety resulted from the lack of paper documentation, difficulty using the alpha code on the return envelope, and general "techno-phobia." In addition, the Southwest DORA's 1997 profile determined some of the barriers to TeleFile participation are created inadvertently by the IRS. For the 1997 filing season, the IRS mailed 26.6 million TeleFile packages to taxpayers. However, 8.2 million of those taxpayers were ineligible. The researchers also found 5.0 million qualified to use TeleFile were not identified or mailed a TeleFile package by the IRS.

These reports and findings provide the IRS with many insights to the barriers to electronic filing growth in the 1990s. They also point to the need for independent marketing strategies for the distinct e-file markets – non On-Line (i.e., preparer-originated), On-Line, and TeleFile.

### **Breaking the Barriers – Marketing e-file**

#### ERO/Practitioner Standard e-filing

The North Florida and Kentucky-Tennessee DORAs' *Alternative Ways of Filing Research Report* (1997) presented research that tested the effectiveness of different treatments directed at electronic return originator participation in the electronic filing program. These treatments included various forms of direct contact with the EROs designed to encourage greater participation. The treatments failed to motivate the EROs, who cited costs and lack of client demand as impediments. The researchers therefore recommended marketing e-file directly to taxpayers, who in turn would create demand for e-file services from their preparers.

Not coincidentally, in 1997 the ETA began developing marketing plans directed at both the tax practitioner community and taxpayers that used practitioners. First, the ETA re-emphasized its marketing to practitioners, highlighting strides toward complete office automation and reduction of paper, confirmation of receipt, and reduced errors. The practitioners had the lucrative RAL (and RAL-like) markets as added incentive. Second, the IRS marketed directly to taxpayers through print advertising, as well as selected radio and television spots, promoting faster refunds and suggesting they ask their preparers about e-filing options.

### On-Line Filing

The North Florida and Kentucky-Tennessee DORAs' recommendations for the ERO/practitioner e-file marketing apply to the On-Line market, as well. A large portion of the IRS's marketing promotes e-filing in general terms, whether it is through a practitioner, on-line, or by telephone. However, the IRS needs to develop a more directed effort towards the On-Line market segment. Among suggestions made by the participants of the 1997 On-Line filing focus groups conducted by the Strategic Planning Division: the IRS should run a marketing campaign simply to inform the public about On-Line filing. Given the participants' misperceptions and anxiety, this seems a good idea. During the 1998 and 1999 filing seasons, the IRS partnered with major tax preparation software companies to advertise On-Line filing capabilities as a selling point for the software.

### TeleFile

In 1996, the IRS contracted Price Waterhouse to develop a marketing and communications strategy to promote e-file use. In its report, *Electronic Filing Marketing and Communication Plan* (1997), the consultants concluded TeleFile was the only IRS electronic filing product ready for full marketing. It found 99 percent of all TeleFile users in 1996 planned to use TeleFile again the following year. However, citing the 1996 *Automated Survey of TeleFile Users*, it also found 85 percent of those eligible to use TeleFile did not do so. These two facts pointed to a marketing deficiency.

The DORAs explored this deficiency. Citing that only 11 percent of all taxpayers receiving

TeleFile packages used TeleFile in 1996, the Connecticut-Rhode Island DORA conducted local research aimed at increasing this rate. Its *Research Report: The Effect of a Mailing on TeleFile Usage* (1998) describes the test it conducted to measure the effect of flyer distribution on TeleFile usage in 1997. Though limited in scope, the researchers concluded mailing reminder flyers in addition to TeleFile packages has no effect on participation. In 1997, the North Florida and Kentucky-Tennessee DORAs conducted a similar test in the IRS Southeast Region, using postcards rather than flyers. Their *Alternative Ways of Filing Research Report* (1997) arrived at the same conclusion: a separate reminder mail out has no effect on TeleFile participation. These tests confirmed the 1994 TeleFile Customer Satisfaction Survey and Bonney & Co. survey (1996) findings that taxpayer awareness of TeleFile results from the receipt of the TeleFile package itself.

In its *TeleFile Marketing Strategies for the Georgia District* (1998), the Georgia DORA tested three low- or no-cost TeleFile promotion strategies: 1) community and student cable television messages, 2) an advertising campaign implemented one week before the TeleFile package mail out, and 3) staffers in on-campus student mailboxes. The research found the combination of the first and second strategies increases TeleFile participation the most.

However, the researchers acknowledge lack of control on various aspects of the test. They also conclude there may be a natural 25 percent participation rate ceiling, given the programs current restrictions.

The Georgia DORA issued another report, *Expanding TeleFile Eligibility* (1998), that examined the impact of removing various TeleFile eligibility conditions. The research found under 21 percent of Georgia district taxpayers were eligible to use TeleFile in 1997. By removing the taxable income, interest income, and filing status limitations, as well as the restrictions based on age and blindness, eligibility rose to over 32 percent. If in addition, the Service were to allow TeleFile filers to claim two dependents and file a Schedule A for itemized deductions, eligibility would increase to almost 48 percent. The report acknowledges there are

costs associated with lifting the current restrictions, and recommends further study.

## Conclusion

The IRS works with a variety of stakeholders in an effort to expand electronic information exchange. A vital step is the identification and removal of barriers. Barriers exist relative to the IRS's tax administration duties, and to the public's acceptance and use of new technologies. From the public's perspective, the IRS will have to attract electronic exchange by reducing the public's burden and costs in dealing with the IRS. The IRS also must address such issues (real or as perceived by the public) as fraud prevention, electronic authentication, taxpayer privacy and information security, rules for certifying electronic tax administration participants, and how to regulate third parties that wish to exchange information.

This article provides a review of the research done in the individual e-file area. Further work should include annual maintenance of the databases developed to study the individual e-file market, as recommended by the Southwest DORA and others. Also, more research-oriented work is needed in the testing of marketing treatments (e.g., testing different advertising approaches to determine which work best, through the use of test and control groups) and product development. To meet the 80 percent participation goal set by the Congress, the IRS is preparing to implement many new initiatives intended to expand e-file. The IRS should precede these initiatives with small-scale (i.e., low cost) tests, such as those conducted by the Georgia DORA for TeleFile, that quantify their impact. After such tests, the IRS can choose and implement the most effective initiatives.

Of equal importance is the need to start basic research in the business return e-file area. The IRS currently accepts Form 1041, U.S. Fiduciary Income Tax Return (for estates and trusts), Form 1065, U.S. Partnership of Income, Schedules K-1 and series 5500 returns, Annual Return Report of Employee Benefit Plan, on either magnetic tape or via electronic filing. It accepts Form 941, Employers Quarterly Federal Tax Return via magnetic tape, electronic filing, and TeleFile, as well as Form 940 via magnetic tape. These electronic commerce programs

(excluding the magnetic media tape) are new, and still in the developmental stages in terms of the ETA's marketing initiatives. Marketing initiatives and further product development require the support provided by the profiling, study, and treatment testing that have characterized the individual e-file research.

Finally, the Price Waterhouse report, *Electronic Filing Marketing and Communication Plan* (1997), recommends the IRS fully integrate market research into e-file product development from the start to insure products and services are developed with the customer in mind. It goes on to recommend the IRS institutionalize customer feedback into all product development and refinement activities. Such information is vital to planning effective marketing strategies.

### Research Currently in Progress

The IRS's Office of Research and District Offices Research and Analysis each year present IRS management a Research Plan that lists significant research projects planned and underway.<sup>48</sup> The ETA Research Strategy Projects listed on the fiscal year 1999 Research Plan reflect a marked increase in e-file research activity. Notably, in the business return area, Project 1.02 supports the development of an e-file database that will have the same purpose as the individual return e-file database created and used by the Southwest DORA. The Research Plan also contains projects that support analysis of recent ETA pilots and initiatives designed to further expand individual e-file participation. Projects 1.11 and 1.12 entail survey research to determine the impact and success of the alternative signature pilots conducted during the 1999 filing season for the On-Line, as well as preparer/ERO, filed returns. Project 1.05 will continue the exploration of CHAID to project the cost savings of individual e-file returns. There are a total of ten projects in the FY 1999 Plan designed to better understand and promote e-filing.

<sup>48</sup> For IRS employees, the FY 1999 Research Plan is available at <http://www.research.irs.gov/PLANS/plans.htm>.

---

### About the Author(s):

**Javier Framinan** is an economist in the Projections and Compliance Studies Group, National Office of Research. He received an M.A. degree in economics in 1987 from the University of Virginia. He has been with the IRS since 1987

## References

- Angell & Co. (1998). *IRS PSA Tracking Study 'Benchmark' Wave Table Report*. Washington, DC (?)
- Bonney & Co. (1996). *Findings of Focus Groups Regarding the TeleFile Program*. Virginia Beach, VA.
- Bonney & Co. (1996). *Results of Surveys Among Persons Who Filed their 1995 Personal Federal Income Taxes Using the TeleFile Program*. Virginia Beach, VA.
- General Accounting Office. (1993). *Opportunities to Increase the Use of Electronic Filing*. (GAO/GGD-93-40). Washington, DC
- General Accounting Office. (1995). *Electronic Filing Falling Short of Expectations*. (GAO/GGD-96-12). Washington, DC.
- Internal Revenue Service. (1991). Musselman, Bryan. *Electronic Filing – Who's Participating and Who Isn't*. Publication 1500, The IRS Research Bulletin. Planning and Research. Washington, DC.
- Internal Revenue Service. (1994). *1994 TeleFile Customer Satisfaction Survey*. Statistics of Income. Washington, DC.
- Internal Revenue Service. (1996). *Incentives and Disincentives for Electronic Filing: A Survey of Studies and Reports to Date*. Compliance Research. Washington, DC.
- Internal Revenue Service. (1996). *Profile Report: Current and Potential Markets for Electronic Filing*. ELF Profiling Project Team. Washington, DC.
- Internal Revenue Service. (1996). Steuer, Anne and Benson, Michelle. *Exploring Current and Potential Markets for Electronic Filing (ELF) Using Tree-Structured Analysis*. North Central District. St. Paul, MN.
- Internal Revenue Service. (1997). *A National Profile of IRS e-file Users in 1997*. District Office Research and Analysis, Southwest District. Phoenix, AZ.
- Internal Revenue Service. (1997). *Critical Issues for Development of an Electronic Tax Administration Strategy: A Plan for Moving Federal Tax Administration into the Information Age*. Electronic Tax Administration. Washington, DC.
- Internal Revenue Service. (1997). *Identification of Barriers to On-Line Filing Study*. District Office Research and Analysis, Pennsylvania District. Philadelphia, PA.
- Internal Revenue Service. (1997). *On-Line Filing Program: Focus Group Report*. Opinion Research. Washington, DC.
- Internal Revenue Service. (1997). *Alternative Ways of Filing Research Report*. District Offices Research and Analysis, North Florida and Kentucky-Tennessee Districts. Jacksonville, FL; Nashville, TN.
- Internal Revenue Service. (1998). *FY 1996 Cost of Submission Processing in the Service Centers*. Document 10806 (03-98). Office of Cost Analysis. Washington, DC.
- Internal Revenue Service. (1998). *A Strategy for Growth*. Publication 3187. Electronic Tax Administration. Washington, DC.
- Internal Revenue Service. (1998). *The Effect of a Mailing on TeleFile Usage*. District Office Research and Analysis, Connecticut-Rhode Island District. Hartford, CT.
- Internal Revenue Service. (1998). *TeleFile Marketing Strategies for the Georgia District*. District Office Research and Analysis, Georgia District. Atlanta, GA.
- Internal Revenue Service. (1998). *Expanding TeleFile Eligibility*. District Office Research and Analysis, Georgia District. Atlanta, GA.
- Internal Revenue Service. (1999). *Fiscal Year 1999 Research Plan*. Office of Research. Washington, DC.
- Internal Revenue Service. (1999). *On-Line Filers: ETA Market Research for 1998*. District Office Research and Analysis, Southwest District. Phoenix, AZ.
- Klemm analysis Group, Inc. (1996). *'On-Line' Filing Program Customer Satisfaction Survey*. Tir-95-0038. Washington, DC.
- Nelco, Inc. (1994). *Why Tax Preparers Do Not Offer Electronic Filing*. Green Bay, WI.
- Pacific Consulting Group. (1996). *Minimizing Taxpayer Burden and Reducing Costs at the IRS: Analysis of Customer Experience with On-Line Filing*. Palo Alto, CA.
- Price Waterhouse LLP. (1997). *Electronic Filing Marketing and Communication Plan*. Tir-95-0037. Bethesda, MD.
- Tax Systems Modernization Institute, IIT Research Institute. (1997). *Tax Administration Electronic Commerce Strategic Implementation Plan*. Tir-93-C-00026. Lanham, MD.

## Inadequate Compensation to 1120S Corporate Officers Study Report

By **Bruce Korbesmeyer**  
Kansas-Missouri District Office  
Research and Analysis  
June 1998

*Research indicated that nationwide almost \$284 million of additional employment tax was due because officers of small corporations underreported their compensation.*

The Kansas-Missouri and Ohio District Offices Research and Analysis (DORAs) conducted a study of Form 1120S Officer Compensation. The objectives were: (1) to quantify the amount of underreported officer compensation and the associated employment tax gap, and (2) to identify the underlying causes of underreporting and the tax gap associated with each cause. Using the tax year (TY) 1994 Business Returns Transaction File (BRTF) to frame the population, a random sample of 528 TY 1995 returns was chosen for classification. We selected 141 returns for audit.

The study results indicated that nationwide almost \$284 million of additional employment tax was due because officers underreported their compensation. Major causes of underreporting and the associated tax gap for each cause were:

- Intentional Underreporting                      \$106,800,000
  
- Misinterpretation of Reporting Requirements    \$ 32,500,000
  
- Taxpayer's Lack of Tax Law Knowledge            \$ 76,300,000

(All estimates use the lower confidence limit. The total does not add up to \$284,000,000 due to additional sampling error that occurs when estimating the projected employment tax gap by cause.)

The study proposed: (1) the DORAs partner with stakeholders of the potential treatments in the next phase of the research process, and (2) potential treatments be refined through a detailed cost/benefit analysis.

## Can Demographic Trends Predict Taxpayer Noncompliance?

**By Kim M. Bloomquist**  
 Illinois District Office  
 Research and Analysis  
 June 1998

*The Bureau of Economic Analysis' wage and salary adjusted gross income (AGI) "gap" has exhibited a statistically significant relationship with the percent of US population between 35 and 54 years of age. This relationship is used to predict the direction of taxpayer noncompliance to the year 2020.*

This analysis examined how would the aging of the "baby boom" generation influence individuals' compliance with tax laws. The paper introduced a life-cycle view of taxpayer noncompliance that associated noncompliance behavior with age-related factors such as an individual's knowledge of tax issues, motivation to comply and opportunity to underreport earnings. This study also compared IRS and Bureau of Economic Analysis (BEA) estimates of misreported wage and salary income and recommended use of the BEA's Adjusted Gross Income (AGI) Gap for wages and salaries as an indicator of the direction, but not the magnitude, of noncompliance. Regression analysis found that the percent of US population between the ages of 35 and 54 accounted for 84 percent of the variance in the relative wage and salary AGI gap from 1947 to 1995.

Based on these empirical findings, individual noncompliance was projected to increase until the turn of the century as the last of the baby boom generation entered middle age. This trend is expected to reverse early in the next century and aggregate compliance gradually improve as baby boomers leave their peak earning years and begin to retire.

The paper noted several important policy implications for compliance research. First, a prior IRS goal of a three percent reduction in the "tax gap" by the year 2001 is made more difficult because of demographic trends that augur more reporting noncompliance, not less, at least until the turn of the century. Second, IRS researchers

will need to net out the influence of an aging population on aggregate compliance behavior to correctly evaluate the effectiveness of proposed treatment programs. Third, IRS should explore ways to motivate younger taxpayers to comply. Wholesale

education and treatment efforts focusing on younger taxpayers may potentially pay for themselves many times over for a relatively modest cost. Finally, methodology used by the BEA to estimate the AGI Gap points to the need for an industry by industry review of reporting compliance behavior by employers. Major structural changes in the US economy over the last twenty years could mean that a significant portion of employee income is not being reported to the IRS by employers.

## Pre-Payment Position and Income Tax Non-Compliance

By Peter D. Adelsheim, Ph. D.,  
Pacific Northwest District Office  
Research and Analysis  
January 1997

*This research examined the relationship between income tax withholding and compliance.*

Although only 26 percent of all individual returns are under-withheld, they account for over 69 percent of the understated tax. Furthermore, the under-withheld returns have an average tax change of \$820 compared to only \$128 for the over-withheld. This study sought to answer two questions. First, can under-withholding be regarded as a cause of noncompliance? That is, if withholding procedures were more accurate (so as to reduce the number of under-withheld returns), would we expect an improvement in compliance? Second, what is the compliance impact of under-withholding? What incremental revenue might be expected from making particular changes to the system?

Our conclusions were as follows. First, the under-withholding phenomenon was real: the under-withheld were less compliant. Second, this relation was not merely statistical. While the current analysis provided ample evidence that withholding is an

important causal agent in the compliance process. Third, changes to the withholding system could be devised that would have significant impact on revenues at little real cost to the government.

Further study should focus on the following questions. First, are there important, identifiable sub-segments among the 27 million under-withheld returns that are relevant to tax administration? Can research identify groups of taxpayers that are homogeneous with respect to the causes of their under-withholding? Second, although this report focused almost entirely on reporting accuracy, it is important to know whether under-withholding is related to other forms of non-compliance (i.e., filing

and payment). Third, can the alternative withholding systems suggested here and other places be more completely defined? Can estimates of the costs and benefits of these systems be developed?

## Farm Labor Contractors Compliance with the Income Tax Laws

By **James L. Zanetti and  
Marlene Le**

Central California District Office  
Research and Analysis  
January 1999

*Farm labor contractors in California have a generally low level of compliance with the income tax laws when compared with other segments the IRS has studied.*

The United States is the largest agriculture producer in the world, and the State of California is the largest agriculture producer in the United States. According to the 1992 Census of Agriculture, California had over 77,000 farmers on 28.9 million acres growing products with a market value of \$17.0 billion. Wage and salary employees in agriculture accounted for 361,000 (or 2.3%) of California's civilian labor force of 15.3 million. The wages paid to grow and harvest the State's agriculture products were \$2.9 billion.

Much of the labor in the agricultural segment is provided by farm labor contractors (FLCs). The agricultural laborers are employed directly by the FLCs, and subcontracted to farm enterprises. In order to measure FLCs' compliance with the income tax laws, we obtained a database of all of

them registered in Arizona, California, Nevada, and New Mexico from the U.S. Department of Labor (DOL). IRS data for all registered farm labor contractors came from the Individual Returns Transaction File (IRTF) for the calendar years ending December 31, 1992 and December 31, 1993. Secondary data sources included the 1992 Census of Agriculture, the Bureau of Labor Standards Consumer Expenditure Survey, and the California Statistical Abstract - 1994.

The overall U.S. population's level of compliance with the income tax laws is in the mid-eighty percentile. The following table shows comparatively lower compliance levels for California farm labor contractors. These results were used to support local compliance projects in this market segment.

<b>Income Tax Year</b>	<b>Total Number of FLCs</b>	<b>Total Income Tax Reported</b>	<b>Estimated Unreported Income Tax</b>	<b>Average Unreported Tax</b>	<b>Voluntary Compliance Level</b>
<b>1994</b>	3,588	\$8,670,268	\$3,808,317	\$1,061	69%
<b>1993</b>	3,519	\$7,258,998	\$3,460,331	\$983	68%

## Erroneous PBA/PIA Profile

**By Stan Griffin**  
Central California District Office  
Research and Analysis  
January 1999

*Analysis suggests invalid PBA/PIA codes are not an important factor in taxpayer compliance.*

Prior IRS studies indicated a significant portion (up to 30 percent) of Principle Business Activity/Principle Industry Activity (PBA/PIA) codes on filed returns are inaccurate. This research focused on the relationship between PBA/PIA code accuracy and taxpayer compliance. Were taxpayers that used invalid PBA/PIA codes less compliant than those that used valid ones?

We profiled Form 1040 (with Schedule C), Form 1120, Form 1120-S and Form 1065 filers using fiscal year 1994 Interim Compliance Research Information system (ICRIS) sample data files from five states (Massachusetts, Maine, New Hampshire, Vermont, and Pennsylvania) plus Central California. Categorizing results by either valid or invalid PBA/PIA codes, we measured noncompliance several ways.

Analysis of all the measures suggested invalid PBA/PIA codes were not an important factor in tax noncompliance in the Form 1040 Schedule C, Form 1120, Form 1120-S, or Form 1065 populations. However, returns with invalid PBA/PIA codes tended to be filed a few days later than valid PBA/PIA coded returns. This was especially true for business returns (i.e., Forms 1120, 1120-S, and 1065), which had approximately double the *filing* delinquency rate of business returns with valid PBA/PIA codes. Practically speaking, the difference in timeliness for valid versus invalid PBA/PIA coded returns in this instance was approximately one week or less -- not a significant factor in terms of compliance.

All the districts studied reported on average a 5 percent higher rate of invalid codes on Form 1040 Schedule C returns prepared by a paid tax preparer. However, we could not demonstrate a causal relationship between paid tax preparer returns, invalid PBA/PIA codes, and noncompliance.

## Business Profitability and Income Tax Compliance

By **Kim M. Bloomquist**  
 Illinois District Office  
 Research and Analysis  
 January 1999

*From 1957 to 1995, sole proprietor net profit margins and reporting non-compliance exhibited a strong negative correlation. This relationship suggested that as long as competitive market forces act to squeeze small business rates of return, sole proprietor tax compliance will continue to erode.*

In 1992, the IRS estimated the “tax gap”, the amount of tax legally owed but not voluntarily paid, at \$95.3 billion. Of this amount, \$39.9 billion, or 42 percent, was due to unreported small business income. The high rate of noncompliance by small business is often attributed to the lack of third party reporting of business income. In the absence of enforced withholding, the IRS continues to rely on voluntary reporting by business taxpayers backstopped by traditional enforcement techniques (e.g., audits). A more recent emphasis on “wholesale” compliance measures, coupled with a renewed emphasis on customer service, was also hoped to reduce noncompliance.

However, the rate of small business noncompliance has not been static over time. From 1957 to 1995, the Bureau of Economic Analysis (BEA) estimates the share of sole proprietor income not reported on federal tax returns more than doubled -- from 26% to 53% for non-farm proprietors and from 60% to 125% for farm proprietors (if income not reported exceeds reported income, the ratio can exceed 100%). During the same time, net profit margins shrank an average of 30 percent for a core group of non-farm sole proprietorships and 50 percent for farm proprietors.

The high statistical correlation between net profit margin and percent underreported income (-0.90 for non-farm proprietors and -0.77 for farm proprietors) may indicate some businesses initiate or increase their level of tax evasion activity in response to lean economic times. This finding was similar to that reported by Rice (1992) for small corporations. Assuming a continuation of the historical trend of narrowing profit margins, it was projected that reporting noncompliance of non-farm sole proprietors would rise 36 percent from 1995 levels by the year 2020. Similarly, underreporting among farm proprietors in 2020 was expected to rise 27 percent from 1995 levels. This conclusion stood in contrast to a recent analysis of reporting noncompliance of wage and salary income which foresees a gradual improvement in compliance by wage earners over the next 25 years (Bloomquist, 1998).

The paper recommended that the IRS track industry profit trends to provide possible early warning of industries with emerging compliance issues. Alternative tax systems appropriate for small business, such as the Value Added Tax (VAT), also were suggested for further study as a compliance improvement strategy.

## Tax Table Clustering: Empirical Evidence of Non-compliance?

By Peter D. Adelsheim, Ph. D.,  
Pacific Northwest District Office  
Research and Analysis  
May, 1996

*This research examined the question of whether taxpayer distribution within tax table brackets was evidence of non-compliance and whether this information could be a benefit to the IRS.*

In a 1985 article, Joel Slemrod (University of Michigan) developed the idea of secondary evasion in order to study the compliance impact of several policy variables: most notably, marginal tax rates. Secondary evasion committed by taxpayers finding themselves in the bottom of one tax bracket under-report taxable income just enough to slip into the top of the next lower bracket.

Empirically, Slemrod does find a larger than expected proportion of taxpayers in the upper parts of the tax table brackets. Theoretically, he argues secondary evasion is a sure sign of primary evasion and uses the proportion of taxpayers in the upper quintile of each bracket as an empirical measure of noncompliance (in reporting accuracy).

The Pacific Northwest District Office Research and Analysis (DORA) studied Slemrod's idea of secondary evasion to determine its validity and applicability in the IRS. In particular, we sought to answer the following four questions. *First*, does clustering exist? Using the 1988 Taxpayer Compliance Measurement Program (TCMP) database, we too found a disproportionately large number of taxpayers in the upper quintiles of the tax brackets in virtually every market segment examined. *Second*, is clustering evidence of secondary evasion? Although at the outset it was important to answer this question, we concluded it can not be answered empirically. It is difficult to imagine the research design that would separate primary evasion from secondary. *Third*, is clustering evidence of primary evasion? Using the TCMP database, we

could not find any association between clustering and under-reporting of taxes. *Fourth*, can clustering be used to refine IRS's Discriminate Function (DIF) system of case selection? Unfortunately, qualifying DIF selection by reporting income in the upper quintile of a tax table bracket lowered the expected tax change and increased the estimated level of compliance.

Finally, Charles Christian (Arizona State University) has provided an alternative explanation for clustering. He has suggested taxpayers finding themselves near the bottom of a tax table bracket expend the extra energy to find legitimate deductions to move them into the lower bracket. In summary, we found little to encourage further research into the concept of secondary evasion.

## Empirical Goodness of Fit Scoring System (Applied to Tax Preparer Due Diligence)

**By Curt Hopkins**

Central California District Office

Research and Analysis

January 1999

*The results of an Empirical Goodness of Fit (EGoF) scoring system were compared with the opinions of Examination Division Return Preparer Coordinators. The purpose of this effort was to either support or refute the use of EGoF as a method to classify tax preparers as either 'questionable' or 'not questionable' for research purposes.*

The Internal Revenue Service spends a great deal of resources identifying, penalizing, and prosecuting professional tax return preparers who fail to use due diligence. The Examination Division of each District Office maintains a staff of one or more Return Preparer Coordinators (RPCs) to monitor questionable preparers. This project, which was limited to individual income taxes related to Form 1040, compared the results of an Empirical Goodness of Fit (EGoF) scoring system with the relatively resource-intensive work by Examination Division RPCs.

The premise behind EGoF was that tax preparers cannot make up numbers at random on a client's returns, especially if their goal is to minimize the taxes paid by their clients. To this end, the leftmost (first significant) digit of a line item is observed and compared to all leftmost digits of the same line item on returns filed through all tax preparers in a district. The technique used in scoring a specific preparer's change from the expected proportions of ones, twos, threes, etc. is Pearson's Goodness of Fit. Thus the empirical distribution of digits scored by this goodness of fit technique gave each preparer a single score of relative 'fit' within the preparer community as a whole

The purpose of this initial effort was to either support or refute the use of EGoF as a method to classify tax preparers as either 'questionable' or 'not questionable' for research purposes. This determination was made without reference to the magnitude of tax changes or penalties assessed. It is not intended for use in selecting workload for enforcement purposes.

For this project we gave similar packets of tax preparer client data to recognized experts, the

RPCs, in the examination divisions of four test districts. The 366 responses of the experts were compared to the predictions of the mathematical scoring system, and the comparison analyzed with nonparametric statistical techniques. We found the EGoF scoring technique performed as well as the experts. We also found that at successively higher EGoF scores (indicating a greater likelihood of a questionable preparer), fewer mismatches occurred between Goodness of Fit and experts' conclusions. False positives drop below 3 percent at higher scores. The Empirical Goodness of Fit system passed all of the statistical tests for use in grouping preparers for research use.

## Processing Year 1998 Criminal Investigation Questionable Refund Formula Development

**By James A. Wilhelm**  
Georgia District Office  
Research and Analysis  
April, 1998

*The development of updated Questionable Refund Formulas helped Criminal Investigation identify and stop more erroneously filed refund returns.*

Using a mixture of Criminal Investigation (CI) questionable refund data with Processing Year 1997 (PY1997) individual master file data, the Georgia District Office Research and Analysis (DORA) developed the PY1998 Questionable Refund (QR) formulas. These formulas were designed to help CI detect and “freeze” the refund of filers who have filed erroneous refund returns. The PY1998 QR formulas were an improvement over the operational PY1997 QR formulas. In PY1997, CI reviewed more than 3.2

million returns to find 16,532 erroneously filed returns. By using the developed PY1998 QR formulas, CI was expected to be able to review less than 1.5 million returns to identify about 18,000 erroneously filed returns. Thus, the developed PY1998 QR formulas indicated that CI could detect and “freeze” more erroneously filed returns while reviewing over 50 percent fewer returns.

## Roadmap to the Future: Using Groupware as an Analytical Tool to Rank Emerging Compliance Issues

By Deborah J. Myers, Ph.D.  
North Texas District Office  
Research and Analysis  
January 1998

### *Collaborative software aids in prioritizing emerging compliance issues*

In 1997, the North Texas District culminated a two-year effort to study the District's demographic makeup and economic trends by recommending a series of implementation steps the District could take to address emerging compliance issues. A critical step was identifying and prioritizing the top issues for the District. A non-subjective and quantifiable approach was sought to accomplish this task. Once issues were fully brainstormed and fleshed out, a way to rank the various issues was employed. Decision-making software -- known as groupware -- was used to aid group collaboration and facilitate the prioritization process.

A survey of both implementation team members and upper-level District management helped identify the criteria that would be used to evaluate 29 emerging issues. The criteria were: risk to the IRS organization; impact on the IRS; durability of the issue over time; financial cost of implementation; impact on the taxpayer; and, resource availability. In order to rank the issues based on these weighted criteria, the team used groupware as a computer aid. The advantages of using the groupware tool were threefold: 1) team members' votes on the importance of issues remained anonymous throughout the ranking process; 2) the tool allowed a qualitative, potentially contentious ranking session to be quantified in an objective way; and 3) it saved many hours of meeting time. Using groupware proved to be a valuable part of our effort.

## Closed Case Analysis to Support Collection of Agreed Deficiencies

**By Ronald L. Edgerton**  
Ohio District Office  
Research and Analysis  
May 1997

*A closed case database analysis revealed opportunities to improve Examination Division's collection of agreed assessments.*

The Ohio district Office Research and Analysis (DORA) conducted a study for the Ohio District Examination Division of "Trends in the Collection of Agreed Examination Deficiencies" to supplement Examination's efforts to enhance the percentage of 'agreed' dollars it collects. Utilizing variables in Examination Division's fiscal year (FY) 1996 Closed Case Database (CCDB), the Ohio DORA explored a wide array of characteristics of assessments such as the size of assessment, organizational unit making the assessment, business activity of the taxpayer, case cycle time, and the CCDB "collectibility indicator."

Using a variety of illustrative charts and graphs, the Ohio DORA provided Examination Division with a report that highlighted those areas that provide the best opportunities for improving the collection of agreed deficiencies.

Combining this information with input from managers and employees, Examination Division crafted an action plan involving all Examination functions, to take advantage of the low and no cost actions the study brought to light.

## Final Study Report: Self-Employed Real Estate Agents

**By Rick Denesha**  
Upstate New York District Office  
Research and Analysis  
January 1999

*Reporting noncompliance and payment noncompliance by self-employed real estate agents appeared to be caused by poor internal accounting controls, a lack of knowledge of the rules for substantiating deductions, and the lack of withholding at source. Traditional enforcement programs did not appear to be the most efficient way to manage noncompliance within this type of small business market segment.*

A profile report issued in fiscal 1996 indicated self-employed real estate agents were at risk for filing, payment and reporting accuracy problems relative to their income tax. However, Taxpayer Compliance Measurement Program (TCMP) and Automated Issue Identification System (AIIS) data strongly indicated the major issues associated with reporting noncompliance (i.e., accuracy problems) were not unique to real estate agents, but appeared characteristic of all sole proprietorships. The TCMP data further indicated the causes of reporting noncompliance may be the result of taxpayers' poor internal accounting controls and lack of knowledge of substantiation rules for business deductions.

Analysis of Audit Information Management System closed case data in conjunction with AIIS data indicated traditional audit programs were not the most suitable mechanism for managing and improving reporting compliance within this market segment. Regarding payment, the research suggested noncompliance is associated with a lack of withholding on commission payments. Regarding filing, professional license data may be used to identify potential nonfilers. The research concluded compliance problems for these types of small business market segment were best addressed by a proactive, wholesale approach rather than one-on-one enforcement.

## Subchapter C Corporations with Low Retained Earnings

**By Michael Hill**

Central California District Office  
Research and Analysis

**Richard Denesha**

Upstate New York District Office  
Research and Analysis

June 1997

*The Central California and Upstate New York DORAs found corporations with a history of losses were at risk for filing and payment noncompliance, and understatement of tax liability was a function of return complexity.*

Baselining of corporate tax reporting revealed corporations that reported little retained earnings on Schedule L of their Form 1120 had a Voluntary Reporting Percentage (VRP) materially lower than the VRP for other corporations. The Central California and Upstate New York District Offices Research and Analysis (DORAs) conducted a national profile to find whether corporations in this market segment were at risk for compliance problems relative to filing, payment, or accuracy. We also wanted to determine whether these corporations constituted an identifiable and potentially reachable market segment that merited future research.

We used the Interim Compliance Research Information System (ICRIS) databases to analyze the Central California, Upstate New York and National market segments. These data files consisted of a random sample of corporate tax returns processed during 1994. We found the following.

- Over sixty percent of all Internal Revenue Code subchapter C corporations have low retained earnings.
- The majority of these corporations are closely held.
- Over 80 percent report \$500,000 or less in total income.
- Approximately 90 percent report the book value of their assets to be \$250,000 or less.
- Depending on their geographic location, 40 to 50 percent classify themselves as service or retail establishments.
- The effective tax rate for over 95 percent of these corporations was 5 percent or less.
- More than half of these corporations pays no Federal Income tax.

From our profile we concluded (a) corporations with a history of losses, as evidenced by negative retained earnings, are at risk for filing and payment problems; and (b) accuracy issues and understatement of corporate income tax liability were a function of the complexity of the return. However, one can not completely understand the compliance risks associated with a corporation with low retained earnings without considering the other tax returns associated with that corporation. These would include other returns filed by the corporation, as well as the returns of related persons. Finally, (C) corporations with low retained earnings do not constitute a unique, reachable market segment. The compliance problems we identified were common to all small businesses. These include complying with complex tax laws and the difficulties associated with properly accounting for transactions among related parties and entities.

## Estimated Collections on FY 1998 Accounts Receivable Dollar Inventory

**By Todd Headrick**

Research and Statistics of Income  
Office of Planning and Finance  
Projections and Compliance Studies Group  
*March 1999*

*Estimates indicate that of the \$246.1 billion in gross FY 1998 ARDI, the IRS can expect to collect \$41.4 billion over the next ten years, or just under 17 percent. This implies an ADA for gross FY 1998 ARDI of about 83 percent.*

At the request of the Office of Accounts Receivable Analysis, Projections and Compliance Studies Group staff developed estimates of collections on fiscal year (FY) 1998 accounts receivable dollar inventory (ARDI).

These estimates of future collections were derived from historical experience. They can be used to compute an associated allowance for doubtful accounts (ADA) for current and future ARDI levels. The data and extrapolations were grouped by source of assessment and by age of assessment. The estimates indicated that of the \$246.1 billion in gross FY 1998 ARDI, the IRS could expect to collect \$41.4 billion over the next ten years, or just under 17 percent. This implied an ADA for gross FY 1998 ARDI of about 83 percent.

## MF TDA Inventory in Los Angeles District

**By Scott Mendelson**  
Los Angeles District Office  
Research and Analysis  
March 1997

*BMF TDA inventory, primarily comprised of employment tax accounts, declined unexpectedly compared to IMF TDA inventory in Los Angeles District during fiscal years 1996 and 1997.*

This study investigated causes of the substantial change in BMF (Business Master File)/IMF (Individual Master File) Taxpayer Delinquency Account (TDA) inventory mix during fiscal years (FYs) 1996 and 1997. BMF balance due accounts comprised 54 percent of the total TDA inventory during the first quarter of FY 1996. However, IMF TDAs increased by more than 900 percent in January 1996 and by another 600 percent in September 1996. Following these changes, only 39 percent of the total balance due inventory consisted of BMF TDAs.

The research uncovered four reasons for this change. The first relates to the Resource Workload Management System (RWMS) "queue" criteria. Analysis showed the three Western Region districts with the lowest RWMS criteria (one being Los Angeles) had the highest percentages of IMF TDA inventory.

Second, more than 50 percent of Form 941 (employment tax) TDAs in the Automated Collection Site (ACS) for Los Angeles District had been there for at least 6 months, and 36 percent had been in ACS for at least 1 year. Many of these were overdue for transfer to TDA status, thus lowering BMF TDA inventory. Also, a repeat delinquency strategy for certain Form 941 taxpayers had been in place since June 1994, which further impacted the BMF TDA inventory in Los Angeles.

Third, the match of BMF Tax Delinquent Investigations (TDIs) with a database provided by the State of California Employment Development Department (EDD) allowed for a more expedient closure of ACS, queue, and Collection Field function (Cff) TDIs relating to taxpayers that have gone out of

business. This process prevented quite a few BMF TDIs from erroneously evolving into BMF TDAs. A total of 11,385 BMF TDIs were closed during FY 1996 as "final" based on this match, with an additional 7,654 BMF TDIs closed during the first quarter FY 1997.

Finally, the economic downturn of the early 1990s, particularly severe in Southern California, contributed to the reduction in BMF balance due accounts. This was indicated by a downtown office vacancy rate in Los Angeles County that was well above the national average, and a record number of bankruptcy filings, including a Chapter 7 (No Asset) increase of 475 percent in just one year. High city taxes and fees leading to an exodus of businesses from this area, high numbers of business closures, and a smaller number of newly formed businesses also contributed to the reduction in BMF TDAs.

## Business Licensing as Nonenforcement Approach to Increasing Tax Compliance in the Liquor Industry

By Nancy Richman

North-South Carolina District Office  
Research and Analysis

Scott Mendelson

Los Angeles District Office  
Research and Analysis

November 1998

*The South Carolina legislature enacted a bill that requires full compliance with federal and state income taxes prior to granting a new liquor license. Results of this legislation have encouraged efforts by the IRS and the State of California Alcoholic Beverage Control (ABC) to enact a licensing agreement.*

The North-South Carolina IRS District Office Research and Analysis (DORA) researched the effect of licensing legislation on taxpayer compliance. In 1993 South Carolina implemented legislation that requires local and federal tax compliance for business licensing. Since only individuals may apply for liquor licenses in South Carolina, the legislation affects only the compliance of the individual applicant, not the business entity using the license. As an apparent result, from 1992 to 1994, the number of applicants not related to the business using the license increased from 15 percent to 21 percent.

Using a random sample of liquor license applicants in 1992 (pre-legislation) and 1994 (post-legislation), researchers found the percentage of license applicants who failed to file Individual Income Tax returns by late 1996 declined from 12.6 percent for the 1992 applicants to 8.8 percent for the 1994 applicants. Late, or overdue, filing declined from 14.4 percent for 1992 to 12.1 percent for 1994 applicants. The percentage of license applicants who did not pay their income tax by the due date declined from 10 percent for 1992 to 5.5 percent for 1994 applicants.

None of these compliance measurements showed a statistically significant improvement at the 95 percent confidence level between the pre- and post-legislation years. However, liquor licensees in South Carolina were relatively compliant, compared to other market segments, even before enactment of the

legislation. Lack of significant improvement in compliance also may be attributable to the fact some license applicants may not have been aware of the fairly new law, and perhaps applicants had more time to file their 1992 returns before the filing measurement was taken for the purposes of the study.

In related research, the Los Angeles and Southern California DORAs profiled the liquor industry in a six county area of Southern California. The primary purpose was to identify the types and levels of compliance within the population of liquor licensees, which in turn could be used to help IRS and the State of California Alcoholic Beverage Control (ABC) consider a licensing agreement or other strategy. Findings indicated there is room for improvement in the area of compliance. The percentage of liquor licensees who have filed all required Tax Year 1994 income, partnership and employment tax returns is 83.7%, while only 80.0% have done so timely. The percentage of licensees who have *filed timely* and *paid timely* (all required returns) and have made timely Federal Tax Deposits is only 58.4%.

One of the categories for which the State of California ABC suspends or revokes liquor licenses is "moral turpitude." This currently does not include a requirement for a licensee to remain compliant with tax laws and regulations. IRS has begun using the results of this research to suggest states make Federal tax compliance a condition for securing or renewing a liquor license.

## Analysis of Toll-Free Telephone Demand

**By Jeff Butler**

Research and Statistics of Income  
Office of Planning and Finance  
Projections and Forecasting Group  
November 1997

***A comprehensive analysis of toll-free telephone demand led to recommended changes that increased national access rates from 62 percent in 1997 to 91 percent in 1999.***

At the request of the Executive Officer for Customer Service, a comprehensive analysis of toll-free telephone Level Of Access (LOA) was developed for the 1040, 8815, and 4262 program areas—tax law, notices, and refunds, respectively. Using detailed call volume data from 1996 and 1997, this study examined: 1) the impact of using new LOA criteria in 1998 on toll-free performance and tracking measures; 2) the dynamics of access rates within a peak/off-peak framework; and 3) the interrelationships between access rate factors such as overflows, abandons, and calls answered. Combining the results from these areas ultimately led to a quantitative analysis of how to raise the Level of Access.

For example, an aggregate model combining data for all three program areas showed that for every 1 percent increase in the percentage of calls answered, the abandon rate drops 0.2 percent, the ratio of overflows to abandons drops 0.15 percent, and the LOA increases by 1.1 percent. For every 1 percent decrease in the abandon rate, the LOA increases 6.8 percent; and for each one unit decrease in the ratio of overflows to abandons, the LOA increases by almost 10 percent. A more detailed analysis along these lines was conducted for each program area separately in a peak/off-peak framework. Isolating factors in each time period either increased or decreased the LOA.

Using this approach, an average annual LOA could be computed for a given percentage of calls answered. In the 8815 program area, for

example, the models showed that 65 percent of all calls needed to be answered for an average annual LOA of 80 percent. Since only 34.4 percent of calls were answered in 1997, the implication was a needed increase in staffing proportional to the increase in calls answered to reach a 71 percent LOA. (Adjustments were made for the fact that this relationship was nonlinear.) Scenarios were developed for target LOAs of 75, 80, and 85 percent, and recommendations made for increases or reallocations in staffing by time period. These recommended changes were put into place in FY 1998, during which each program saw a dramatic improvement. For the three program areas combined, the IRS increased its LOA from 62 percent in 1997 to over 90 percent the following year.

## Taxpayer Service Walk-in Study

**By Curtis R. Darling**  
Ohio District Office  
Research and Analysis  
October 1996

*An Ohio DORA and Taxpayer Service study of walk-in customers revealed taxpayers continue to look to walk-in sites as a source of tax forms, and promotion of VITA and TCE services could reduce substantially the demand for return preparation services at walk-in sit*

The Ohio District Office Research and Analysis (DORA) and Taxpayer Service Division conducted a study to secure information regarding walk-in customers. Taxpayer assistors recorded the type of assistance provided and the Zip Code of the taxpayer on log sheets for nine weeks – six weeks during the 1996 filing period and three following the filing period. The study included nearly 50,000 filing period contacts (30 percent of all 1996 filing period contacts) and 7,500 post-filing period contacts (12 percent of all 1996 post-filing period contacts). In addition, information from a survey initiated by Internal Audit at walk-in sites in the Northeast Region, including six sites in Ohio, was included in the study.

The study found 51 percent of all contacts during the filing period were for tax forms and 30 percent were for return preparation services. The percent of contacts for these services after the filing period dropped to 23 percent and 10 percent, respectively. Profiles for each walk-in site provided increased insight into why taxpayers visit particular sites. Contacts regarding forms distribution, return preparation, notices, lien clearances, individual income tax and business-related tax questions, filing issues, highway use tax and total visits were displayed in a collection of tables, colored graphs, and maps that allowed IRS management to identify what taxpayer population were served at each site.

Based on this information, locations were identified where an increase in Volunteer Income

Tax Assistance (VITA) and Tax Counseling for the Elderly (TCE) services could reduce the demand for return preparation services. Sites also were identified which would be ideal locations to develop or improve alternative methods for tax form distribution, thereby reducing the demand for walk-in services related to tax forms. A self-help kiosk to distribute tax forms was one of the possible applications at those sites that have a high demand for tax forms.

## Selected Estimates of Returns with Form 1099 Information

**By Terry Manzi**

Research and Statistics of Income  
Office of Planning and Finance  
Projections and Compliance Studies Group  
June, 1999

*This analysis estimated which types of information documents were prevalent among individual returns filed on or before February 15.*

The IRS Restructuring and Reform Act of 1998 required that a study be done of the impact of extending the deadline for providing taxpayers with copies of information returns (other than Forms W-2) from January 31 to February 15. Staff from the Office of the Assistant Commissioner (Electronic Tax Administration) were responsible for preparing that report. They, in turn, requested that the Projections and Compliance Studies Group develop a profile of the early return filers in terms of their information return characteristics.

The Projections and Compliance Studies Group analysis provided the following broad profile, in terms of approximate return filing volumes in calendar year 1999:

- number of returns filed by February 15 reflecting strictly wages and salaries (i.e., Form W-2 information, only) -- 16,649,000;
- number of returns filed by February 15 with at least some "Form 1099-type" information (i.e., other than Form W-2) -- 14,807,000
- number of "Form 1099-type" returns filed by February 15 with interest income (Form 1099-INT) -- 10,168,000;
- number of "Form 1099-type" returns filed by February 15 with unemployment compensation or taxable state refund (Form 1099-G) -- 5,006,000;
- number of "Form 1099-type" returns filed by February 15 with home mortgage interest deduction (Form 1098) -- 4,104,000;
- number of "Form 1099-type" returns filed by February 15 with pension/IRA distributions (Form 1099-R) -- 2,937,000;
- number of "Form 1099-type" returns filed by February 15 with taxable social security income (Form 1099-SSA/RRB) -- 1,144,000;
- number of "Form 1099-type" returns filed by February 15 with dividend income (Form 1099-DIV) -- 1,868,000;
- number of "Form 1099-type" returns filed by February 15 with capital gains information (Form 1099-B) -- 1,492,000

## Alternative Signature Methods for Filing Individual Tax Returns

**By T. Scott Shutt,  
Robert A. Kerr and  
Dennis L. Raup**

Research and Statistics of Income  
Artificial Intelligence Lab 2  
November 1998

*Both Voice and Digitized Signature methodologies are viable as alternative signature methods to create a totally paperless filing experience.*

Beginning in 1995 two alternative signature methods were tested to provide information as to using voice or a dynamic digitized signature to make electronically filing individual income tax returns totally paperless. The impetus behind these efforts was to (1) broaden electronic filing opportunities, (2) eliminate the required paper signature document which is costly to process and unwieldy to manage, and (3) provide a forensically sound electronic signature as legally binding as a pen to paper signature.

The voice signature test required taxpayers, once their returns were filed electronically, to initiate a telephone call, state their name and identifying information, and read the affirmation text from the return. The affirmation text and the name constituted the signature. All signatures were tested for cooperation, identification and potential fraud (as an impostor or a multiple filer) using speaker verification and recognition techniques. Test forgeries were introduced into the test population for fraud tests.

Focus groups with the test participants indicated little aversion to using voice as a signature method. Results of operations-oriented research indicated the use of voice was technically feasible as an alternative signature method; however the large size of the stored voice signature would have a significant impact on storage requirements if implemented. Speaker verification easily detected cooperativeness. Speaker recognition easily determined impostors

during the fraud tests but had more difficulty with the multiple filers. Also, the test indicated multiple samples of the same person's voice over several years could become a fraud detection tool that could provide indications of suspicious returns quickly. The use of voice as an alternative signature method is viable but needs more study. Voice is a no cost solution to the taxpayer and practitioner in creating a totally electronic return. Voice could provide additional fraud detection and prevention tools.

The Dynamic ELF Signatures Test (DigEST) used routines embedded in tax preparation software to capture the taxpayers' signature from a digitizing pad at the time of tax return preparation.

The major conclusions from the DigEST were (1) the digitized signature technology works and provides more security than a paper signature, (2) the digitized signatures are not as strong forensically as paper signatures but are as strong as any electronic method in a criminal trial, (3) the public accepts digitized signatures when presented positively, (4) IRS can process digitized signatures successfully, at much lower costs than paper signature documents, and (5) practitioners must perceive a business reason for using any electronic signature or they will not use it.

The test demonstrated DigEST is a workable signature methodology for totally paperless tax filing.

## Focus Group Report: On-Line Filing Program

By Dru DeLong and  
David Browne

Management and Administration  
Strategic Planning Division  
July 1997

*Focus group research revealed taxpayers' perceptions of benefits and barriers to individual tax return On-Line Filing.*

Trained moderators from the Strategic Planning Division conducted six focus group interviews with taxpayers eligible to participate in the On-Line Filing Program but did not. These focus groups were conducted at the request of the Pennsylvania Taxpayer Education/Electronic Filing Staff.

Roughly half the participants knew filing income tax returns using home computers was possible, but knew no detailed information beyond that. Most expressed a willingness to use this alternative method of filing. They believed transmitting to IRS directly rather than through a third party would remove most of the barriers to participating in this program. Researchers also found the following.

- Two barriers were common to most of the participants: mistrust of data security and the yearly cost of tax preparation software.
- Publicizing information about the service provider, explaining security measures, and establishing a procedure of accountability could alleviate the data security barrier.
- A clear benefit over filing a paper return is receiving notification that IRS received the return. Other benefits included record storage on disk rather than paper, and receiving refunds faster.
- To be most effective, a marketing campaign should provide information about the on-line Filing Program, how it works, its requirements, and where to get more information.

1. Detail may not add due to rounding.
2. Table 1 shows tax return and economic/demographic data for IRS regions and states.
3. Table 2 shows tax return and economic/demographic data for IRS regions and the new consolidated IRS districts.
4. The years displayed represent the calendar year in which the tax return was filed, except (table 3) Federal tax deposits, which are presented on a fiscal year basis.
5. Economic and demographic data projections were made in July 1999 by Data Resources, Inc., (DRI)/McGraw Hill. This information is not available for the Assistant Commissioner (International). For most recent economic and demographic projections, please direct questions to the contacts listed on the inside cover of this publication.
6. The selected employment by industry variables (e.g., construction employment, mining employment, etc.) do not sum to "Civilian Employment" since they are only a subset of all industries.
7. Federal tax deposit projections were made in March 1999. These figures are organized by the Service Center Recognition Imaging Processing System (SCRIPS) alignments and include projections of electronic fund transfers as specified in the original North American Free Trade Agreement legislation. Withholding and information document projections were made in May 1999 and are also sorted by a SCRIPS grouping of service centers. The tax return projections reflect the fall 1998 editions of IRS Document 6149 (districts and regions) and Document 6186 (service centers). For more recent projections of tax returns, please direct questions to the contacts listed on the inside cover of this publication.
8. Total returns consist of the following tax forms:

<i>Individual</i>	Paper and electronic Forms 1040, 1040A, 1040EZ, 1040PC, 1040NR, 1040PR, and 1040SS
<i>Estimated Tax</i>	Form 1040ES
<i>Fiduciary</i>	Form 1041 and Form 1041ES
<i>Partnership</i>	Form 1065
<i>Corporation</i>	Forms 1120, 1120A, 1120F, 1120FSC, 1120H, 1120POL, 1120REIT, 1120RIC, 1120S, 1120PC, 1120L, and 1120SF
<i>Estate</i>	Forms 706, 706NA, 706GS(D), and 706GS(T)
<i>Gift</i>	Form 709
<i>Employment</i>	Forms 940, 940EZ, 940PR, 941, 941E, 941PR, 941SS, 942, 942PR, 943, 943PR, 945, and CT-1

**Exempt Organization****Forms 990, 990C, 990EZ, 990PF, 990T, 4720 and 5227***Employee Plans**Forms 5500, 5500C, 5500EZ, and 5500R**Excise**Forms 11C, 720, 730, and 2290**Form 8752**Supplemental Documents* *Forms 1040X, 1120X, 2688, 4868, 7004, and 1041(A) prior to 1993*

9. Withholding Documents consist of the following: Forms W2, W2P (prior to 1992) and W2G.
10. Information documents consist of the following: Forms 1098, 1099A, 1099B, 1099DIV, 1099G, 1099INT, 1099MISC, 1099C, 1099OID, 1099PATR, 1099R, 1099S, 1099SSA/RRB, 5498, 1096, Schedules K-1 and foreign information.